



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 131993**

**TO: Celine Qian**  
**Location: REM-2Ad4&2C70**  
**Art Unit: 1636**  
**Monday, September 13, 2004**

**Case Serial Number: 09/939209**

**From: Edward Hart**  
**Location: Biotech-Chem Library**  
**REM-1A55**  
**Phone: 571-272-2512**

**edward.hart@uspto.gov**

### **Search Notes**

Examiner Qian,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

**This Page Blank (uspto)**

Art Unit 1636 Phone Number 302-0777 Serial Number 09/939209  
Mail Box and Bldg. Room Location 2A64 Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*  
Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc., if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: Methods & systems for facilitating the diagnosis & treatment of  
Inventors (please provide full names): Levitt et al. Schizophrenia

Earliest Priority Filing Date 8/24/01

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

*Please search 4110 bp - 12150 bp of SEQ ID NO:3.  
both prior art & interference.*

*3-A9-20300*

*mej*

STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher _____	NA Sequence (#) <u>1</u>	STN _____
Searcher Phone # _____	AA Sequence (#) _____	Dialog _____
Searcher Location _____	Structure (#) _____	Quester Other _____
Date Searcher Reported <u>9/9/04</u>	Bibliographic _____	IP Link _____
Date Completed <u>9/13/04</u>	Litigation _____	Lexis Nexis _____
Searcher Prep & Review Time _____	Fulltext _____	Sequence Systems <u>Q66H</u>
Final Prep Time _____	Patent Family _____	WAXW Internet _____

**This Page Blank (uspto)**



Pending Nucleic Acid and Pending Amino Acid database searches generate two sets of results each. The Pending databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions .rnpn and .rnpn

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions .rapn and .rapn

***Because they contain data that is confidential, the results of Pending database searches should not be left in the case .***

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 11, 2004, 08:12:38 ; Search time 1835 Seconds

(without alignments)  
18615.673 Million cell updates/sec

Title: US-09-939-209A-3\_COPY\_4110\_12150

Perfect score: 8041  
1 caaatggaagcccgggcac.....ctaggatcctcagctgaccc 8041

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: Geneseq1980s:\*  
2: Geneseq1990s:\*  
3: Geneseq2000s:\*  
4: Geneseq2001as:\*  
5: Geneseq2001bs:\*  
6: Geneseq2002s:\*  
7: Geneseq2003as:\*  
8: Geneseq2003bs:\*  
9: Geneseq2003cs:\*  
10: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8041	100.0	20300	6	ABK47337 Genomic n
2	277.6	3.5	1238	4	AA126766 Human bre
3	274.4	3.4	3165	7	ABX63113 Human CDV
4	274.4	3.4	3167	9	ADD12660 Human CDV
5	180.2	2.2	1902	4	AAK94664 Human ful
6	170.2	2.1	11000	6	ABT00010_10
7	170.2	2.1	11000	6	ABT01503_10
8	169.6	2.1	11000	4	AAK95240_10
9	169.6	2.1	11000	4	AAK96733_10
10	168.4	2.1	7457	4	AAK86211 Human imm
11	167.8	2.1	401	4	AAK96142 Human neu
12	167.8	2.1	401	4	AAK97635 Human neu
13	167.8	2.1	401	6	ABT00912 Human neu
14	167.8	2.1	401	6	ABT02405 Human neu
15	166.2	2.1	847	4	AAK91687 Human CDV
16	166.2	2.1	847	4	AAK93880 Human CDV
17	163.2	2.0	3437	6	AB155857 Human GAB
18	163.2	2.0	414	8	ACH30357 Human tes
19	158.2	2.0	11000	6	ABSS5320_2
20	158.2	2.0	34151	6	ABSS5200 Genomic D
21	155.2	1.9	84607	2	AAK90847 Human PAC
22	154.8	1.9	16310	3	AAA34964 Human ace
23	154.8	1.9	16310	3	AAF21086 Human low

24	154.8	1.9	16310	7	ABZ96780 Human nuc
25	154.8	1.9	17634	3	AAA34965 Human ade
26	154.8	1.9	17634	3	AAF21087 Human low
27	154.8	1.9	17634	7	ABZ96781 Human nuc
28	154.8	1.9	81800	6	ABK94756 Human CDV
29	153.8	1.9	11000	6	ABL57909_0
30	150.2	1.9	222880	6	ADCB7622 Human GPC
31	147.6	1.8	126512	6	ABN83429 Human tra
32	147.2	1.8	11000	6	ABSS5320_0
33	147.2	1.8	34151	6	ABSS5200 Genomic tra
34	146.8	1.8	11000	6	ABO74964_1
35	146	1.8	338462	6	ADC24763 Human wil
36	144.8	1.8	326014	6	ABK89296 Human gen
37	144.4	1.8	140167	6	ABT01146 Human bre
38	142.8	1.8	11000	4	AAK95240_03
39	142.8	1.8	11000	4	AAK96733_03
40	142.8	1.8	11000	6	ABT00010_03
41	142.8	1.8	11000	6	ABT01503_03
42	142	1.8	7163	4	AAK46545 Human su
43	141.2	1.8	349881	4	ADC86642 Human GPC
44	141	1.8	41684	3	AAA28150 Human pur
45	140.4	1.7	175590	9	ADD50650 BAC sequ

## ALIGNMENTS

RESULT 1	ABK47337	ABK47337 standard; DNA; 20300 BP.
ID	ABK47337	standard; DNA; 20300 BP.
XX	ABK47337	
AC	ABK47337	
XX	ABK47337	
DT	18-JUN-2002	(first entry)
XX		
DE	Genomic nucleotide sequence encoding human RGS-4 protein.	
XX		
RGS-4	schizophrenia; human; regulator of G protein signalling 4;	
XX		
KW	neuroleptic; polymorphism; schizophrenia; schizotypy; schizoid; chromosome 1q21-22;	
KW	gene; ds.	
XX		
OS	Homo sapiens.	
XX		
XX		
Key		
FT	variation	Location/Qualifiers
FT		replace(4121,T)
FT		/*tag= a
FT		/standard_name= "Single-nucleotide polymorphism"
FT		replace(4123,A)
FT		/*tag= b
FT		/standard_name= "Single-nucleotide polymorphism"
FT		replace(4368,C)
FT		/*tag= c
FT		/standard_name= "Single-nucleotide polymorphism"
FT		replace(4621,C)
FT		/*tag= d
FT		/standard_name= "Single-nucleotide polymorphism"
FT		replace(4790,T)
FT		/*tag= e
FT		/standard_name= "Single-nucleotide polymorphism"
FT		replace(4816,T)
FT		/*tag= f
FT		/standard_name= "Single-nucleotide polymorphism"
FT		replace(4970,T)
FT		/*tag= g
FT		/standard_name= "Single-nucleotide polymorphism"
FT		replace(5055,G)
FT		/*tag= h
FT		/standard_name= "Single-nucleotide polymorphism"
FT		replace(5295,A)
FT		/*tag= i
FT		/standard_name= "Single-nucleotide polymorphism"
FT		replace(5695,A)
FT		/*tag= j

PT	variation	/standard_name= "Single-nucleotide polymorphism"
FT	variation	/replace(17375,T)
FT	variation	/*tag= x
FT	variation	/standard_name= "Single-nucleotide polymorphism"
FT	variation	/replace(17759,A)
FT	variation	/*tag= 1
FT	variation	/standard_name= "Single-nucleotide polymorphism"
FT	variation	/replace(8596,A)
FT	variation	/*tag= m
FT	variation	/standard_name= "Single-nucleotide polymorphism"
FT	variation	/replace(9602, .9610,CA)
FT	variation	/*tag= n
FT	variation	/note= "Deletion polymorphism"
FT	variation	/replace(9892,A)
FT	variation	/*tag= o
FT	variation	/standard_name= "Single-nucleotide polymorphism"
FT	variation	/replace(9565,A)
FT	variation	/*tag= p
FT	variation	/standard_name= "Single-nucleotide polymorphism"
FT	variation	/replace(10132,A)
FT	variation	/*tag= q
FT	variation	/standard_name= "Single-nucleotide polymorphism"
FT	variation	/replace(11056,C)
FT	variation	/*tag= r
FT	variation	/standard_name= "Single-nucleotide polymorphism"
FT	variation	/replace(11091,T)
FT	variation	/*tag= s
FT	variation	/standard_name= "Single-nucleotide polymorphism"
FT	variation	/replace(11106,A)
FT	variation	/*tag= t
FT	variation	/standard_name= "Single-nucleotide polymorphism"
FT	variation	/replace(11774,T)
FT	variation	/*tag= u
FT	variation	/standard_name= "Single-nucleotide polymorphism"
FT	variation	/replace(12143,A)
FT	variation	/*tag= v
FT	variation	/standard_name= "Single-nucleotide polymorphism"
FT	variation	/replace(12145,T)
FT	variation	/*tag= w
FT	variation	/standard_name= "Single-nucleotide polymorphism"
FT	variation	/replace(11367,G)
FT	variation	/*tag= x
FT	variation	/standard_name= "Single-nucleotide polymorphism"
FT	variation	/replace(117027, .117029,GA)
FT	variation	/*tag= y
FT	variation	/note= "Deletion polymorphism"
FT	variation	/replace(17630,T)
FT	variation	/*tag= z
FT	variation	/standard_name= "Single-nucleotide polymorphism"
XX		
PN	W0200216653-A2.	
XX		
PD	28-FEB-2002.	
XX		
FF	24-AUG-2001; 2001WO-US026622.	
XX		
FR	24-AUG-2000; 2000US-0228021P.	
XX		
PA	(UYP1-) UNIV PITTSBURGH.	
XX		
PI	Levitt PR, Mirnice K, Kodavali VC, Nimgaonkar VL;	
XX		
DR	WPI; 2002-292070/33.	
XX		
PT	Diagnosing, assessing susceptibility and treating schizophrenia, involves	
XX	observing regulator of G-protein signaling 4, RGS4 levels in a subject.	
PS	Claim 1; Page 20-33; 112pp; English.	
XX		
CC	This invention relates to a novel method for diagnosing schizophrenia or	
CC	determining susceptibility to schizophrenia in a human. The method	
CC	comprises obtaining from a patient a DNA sample and detecting variations	
CC	in the regulator of G-protein signaling 4 (RGS4) gene. Alternatively,	

Query Match	100.0%	Score 8041	DB 6	Length 20300
Best Local Similarity	100.0%	Prod. No. 0		
Matches 8041	Conservative	0	Mismatches	0
			Indels	0
			Gaps	0
QY 1	CAAAATGGAAGTCTGGGACGCTTACAAAACATAGAGCTCATAAAGTTTACCTTAATTTACC	60		
Db 4110	CAAAATGGAAGTCTGGGACGCTTACAAAACATAGAGCTCATAAAGTTTACCTTAATTTACC	4169		
QY 61	AGATTTCTTAGAGATTTATATTTGAATTCGAGATCTGCTCTCTGTAAAGTTCAACA	120		
Db 4170	AGATTTCTTAGAGATTTATATTTGAATTCGAGATCTGCTCTCTGTAAAGTTCAACA	4229		
QY 121	CTTCCATACACCACTACTGTTTCTTTCACACCTGCACATGCAATGAAATGAACTTTATGAA	180		
Db 4230	CTTCCATACACCACTACTGTTTCTTTCACACCTGCACATGCAATGAAATGAACTTTATGAA	4289		
QY 181	CTGCTGTTTCTATCTCGGGCTAAATAGTTGACAAAAAAGATTATATCTTGGGATAGGC	240		
Db 4290	CTGCTGTTTCTATCTCGGGCTAAATAGTTGACAAAAAAGATTATATCTTGGGATAGGC	4349		
QY 241	TATTTTGGGTTTCTCCACTCTTGTGGAAACAAGGTTTCTTCCCTGGGTAATTAAGT	300		
Db 4350	TATTTTGGGTTTCTCCACTCTTGTGGAAACAAGGTTTCTTCCCTGGGTAATTAAGT	4409		
QY 301	GTGGATATGTTCTTTCACAGGAAATCACTGATGATCACTGCTGTATCAATATGTCAGGG	360		
Db 4410	GTGGATATGTTCTTTCACAGGAAATCACTGATGATCACTGCTGTATCAATATGTCAGGG	4469		
QY 361	TTGGAGTCTCTGATTTATTCATGATGTGCCCAAAAGCTTGTCGAAAGAAATTTGGACACTT	420		
Db 4470	TTGGAGTCTCTGATTTATTCATGATGTGCCCAAAAGCTTGTCGAAAGAAATTTGGACACTT	4529		
QY 421	TCCAAAATAGACATCTGGAAGTCCGTTTACTCTTCGTAATACACATCTCTCC	480		
Db 4530	TCCAAAATAGACATCTGGAAGTCCGTTTACTCTTCGTAATACACATCTCTCC	4589		
QY 481	AGCCCAATCTTTGCTTTTATGTCCTAAAATCAATTAATCTGAATCTGTATGATGCTCA	540		
Db 4590	AGCCCAATCTTTGCTTTTATGTCCTAAAATCAATTAATCTGAATCTGTATGATGCTCA	4649		
QY 541	GCGCATTTGATGAACAAATAAGAAAGGAGGCTTTCGACACTGAGAGGAAATTTGTC	600		
Db 4650	GCGCATTTGATGAACAAATAAGAAAGGAGGCTTTCGACACTGAGAGGAAATTTGTC	4709		
QY 601	ATCTGAAGTGTGCACAGCACAGCTGGGGCTGAGCCTTGACCTCACTCTGCCCCAAGTG	660		
Db 4710	ATCTGAAGTGTGCACAGCACAGCTGGGGCTGAGCCTTGACCTCACTCTGCCCCAAGTG	4769		
QY 661	AGGATATAGTCCCCCATTTTAATCTTGTAATCACTAAAGAACGGAAGCCTGCGCAATGA	720		
Db 4770	AGGATATAGTCCCCCATTTTAATCTTGTAATCACTAAAGAACGGAAGCCTGCGCAATGA	4829		
QY 721	CTTATTTCCCTGCATTTGATACCGTCAATCCTTGAGAAATGTTTCTTTTGTCTCCGTG	780		
Db 4830	CTTATTTCCCTGCATTTGATACCGTCAATCCTTGAGAAATGTTTCTTTTGTCTCCGTG	4889		
QY 781	AGCAAAAGTTGGAAAAATTTGAATTTTACTACAGACACACACTAAGTTCACATCCTGCTG	840		

4890 AGCAAAAGGTGGAAAATTTGAAATTTTAACTAGAGACCAACATAGTTCCATCTCGTGG 4949  
QY TGTGCTGAATGTCTGCCCCCAGTAGAGAAACAGTTCCTTTAAAGCCTATTGTCAACAT 900  
Db TGTGCTGAATGTCTGCCCCCAGTAGAGAAACAGTTCCTTTAAAGCCTATTGTCAACAT 5009  
QY ACCCTCCAGATGTTAGATTTTAACTTAAAGAACTTAAATAGCCTTAAACCTTTTG 960  
Db ACCCTCCAGATGTTAGATTTTAACTTAAAGAACTTAAATAGCCTTAAACCTTTTG 5069  
QY 961 CGAGTTTCTGAAATCCAAATCTATCTTTTACTGCTGCTCCAGAGCTTCTTTAGAA 1020  
Db CGAGTTTCTGAAATCCAAATCTATCTTTTACTGCTGCTCCAGAGCTTCTTTAGAA 5129  
QY 1021 TGCTAACTGATCGGCTTAAGTAACTTAAGTAACTTCTCTCCATTAAGTAACTAAGTA 1080  
Db TGCTAACTGATCGGCTTAAGTAACTTAAGTAACTTCTCTCCATTAAGTAACTAAGTA 5189  
QY 1081 AATTGCTCTTCAAGTAAACAAACACCCCTTAACTTCTAGCATATTCGTGACCTT 1140  
Db AATTGCTCTTCAAGTAAACAAACACCCCTTAACTTCTAGCATATTCGTGACCTT 5249  
QY 1141 CATTTATCTCTCTCTCTCTCTCAAGCTGCAATGGGGTGAACCGTGTGATACATTTTC 1200  
Db CATTTATCTCTCTCTCTCTCTCAAGCTGCAATGGGGTGAACCGTGTGATACATTTTC 5309  
QY 1201 TCTTCATCATTAAGGGTGGCAACCAAACTGCTATAGTAAAGCAAGTTAATAAGACA 1260  
Db TCTTCATCATTAAGGGTGGCAACCAAACTGCTATAGTAAAGCAAGTTAATAAGACA 5369  
QY 1261 AAACCTAAACAATTTATTTAATCAAAAGTTTACATGACATGGAGTCTTCAGAAATGAAG 1320  
Db AAACCTAAACAATTTATTTAATCAAAAGTTTACATGACATGGAGTCTTCAGAAATGAAG 5429  
QY 1321 ACCCAAAAGCCAGGGGAACTGTCTGTTTTTTTGTGAGGTTTCAATGAGAAATGATTA 1380  
Db ACCCAAAAGCCAGGGGAACTGTCTGTTTTTTTGTGAGGTTTCAATGAGAAATGATTA 5489  
QY 1381 GCATGTACCATGTATGATTAACAAAGATATGATCTAGTGAAGAAAGCACTCAGGAGGA 1440  
Db GCATGTACCATGTATGATTAACAAAGATATGATCTAGTGAAGAAAGCACTCAGGAGGA 5549  
QY 1441 AACACAGCAAGGCGCTGTATATTCAGATTTCTTTGATCTCTCTCTCTATGATATGACAT 1500  
Db AACACAGCAAGGCGCTGTATATTCAGATTTCTTTGATCTCTCTCTCTATGATATGACAT 5509  
QY 1501 TCTTCCCTCGAGATGGGCGAGAACTCTTCAATGAGGGTCTCAAGGAGAGAGG 1560  
Db TCTTCCCTCGAGATGGGCGAGAACTCTTCAATGAGGGTCTCAAGGAGAGAGG 5669  
QY 1561 AGAAAGTGGCCTTTTATGATTTTATGCTTGTGCGGGAAGAGAGTCTAGTTCTAT 1620  
Db AGAAAGTGGCCTTTTATGATTTTATGCTTGTGCGGGAAGAGAGTCTAGTTCTAT 5729  
QY 1621 GACCATCTTTGGGGAAGAGAACTCTGTTCTGTGACTTCTTCAATGAGAAAGGA 1680  
Db GACCATCTTTGGGGAAGAGAACTCTGTTCTGTGACTTCTTCAATGAGAAAGGA 5789  
QY 1681 GTTAAAGGACAGAGGCGAGAGATGTTCAGAAAGACCTTGCTGCTTCAAGGCGTCC 1740  
Db GTTAAAGGACAGAGGCGAGAGATGTTCAGAAAGACCTTGCTGCTTCAAGGCGTCC 5849  
QY 1741 GGTCTCTCTTATGTTCCAGTACTTCTTATGCAATACCAAGACATATTTGGGATATGAT 1800  
Db GGTCTCTCTTATGTTCCAGTACTTCTTATGCAATACCAAGACATATTTGGGATATGAT 5909  
QY 1801 TTTCTAGAGCTTAACATGCAATCATGTAAATCTCTCTATGACCTTCAACATTCACAT 1860  
Db TTTCTAGAGCTTAACATGCAATCATGTAAATCTCTCTATGACCTTCAACATTCACAT 5969  
QY 1861 TGCCTTTATCTTTATGATGTGTATGATGCAATAGAGTCAATAGCAAGACCTTGAAGTCCC 1920  
Db TGCCTTTATCTTTATGATGTGTATGATGCAATAGAGTCAATAGCAAGACCTTGAAGTCCC 1920

5970 TGCCTTTATCTTTATGATGTGTATGATGCAATAGAGTCAATAGCAAGACCTTGAAGTCCC 6029  
QY 1921 ACTGTCTGAGCTGGCATTAACATTAATCACTTAATCAATGTGTAAAGTCAAGTAAATGAC 1980  
Db ACTGTCTGAGCTGGCATTAACATTAATCACTTAATCAATGTGTAAAGTCAAGTAAATGAC 6089  
QY 1981 TTAAGTCTCTAAGCTTCAATCTGTAAAGAAATGAAATCATTTGAAGAAACATTCCTCAGAGAT 2040  
Db TTAAGTCTCTAAGCTTCAATCTGTAAAGAAATGAAATCATTTGAAGAAACATTCCTCAGAGAT 6149  
QY 2041 GATCATAGGAATTAAGTAATTAACCTGGCATATAGTCTTAAACAGTCCCTGCTCACT 2100  
Db GATCATAGGAATTAAGTAATTAACCTGGCATATAGTCTTAAACAGTCCCTGCTCACT 6209  
QY 2101 TAGTGAAGATTAATATGCTGTTATTAAGTGGCCACATATGAGTGTGATGCTCTCTCTCT 2160  
Db TAGTGAAGATTAATATGCTGTTATTAAGTGGCCACATATGAGTGTGATGCTCTCTCTCTCT 6269  
QY 2161 TTGTACACGACATACATCTGATTAATCAATTTTAAAGGCTCTCTGTGAAAACCACTGCA 2220  
Db TTGTACACGACATACATCTGATTAATCAATTTTAAAGGCTCTCTGTGAAAACCACTGCA 6329  
QY 2221 GATTCAAAAGATTGAGTTTAACTCTATCTCTGTGCTTCTCTGAGAGTTTGTAAAGTAA 2280  
Db GATTCAAAAGATTGAGTTTAACTCTATCTCTGTGCTTCTCTGAGAGTTTGTAAAGTAA 6389  
QY 2281 ATCTCACTTGAACATCAATGATAGTGTCTTGAAGAACTAACATTCAGTGAAGAGACATA 2340  
Db ATCTCACTTGAACATCAATGATAGTGTCTTGAAGAACTAACATTCAGTGAAGAGACATA 6449  
QY 2341 ACTTAACCAATTTTTTTCTCATCAAGTTATTAATGAATGGCATTATGAATGATGGCA 2400  
Db ACTTAACCAATTTTTTTCTCATCAAGTTATTAATGAATGGCATTATGAATGATGGCA 6509  
QY 2401 TTCAAGGACCTGCTGATCTGTTTCACTTAAGTCACTGTTTCCAAATCATGATGATG 2460  
Db TTCAAGGACCTGCTGATCTGTTTCACTTAAGTCACTGTTTCCAAATCATGATGATG 6569  
QY 2461 ACATTTGAAGACTTAT 2520  
Db ACATTTGAAGACTTAT 6629  
QY 2521 CTGCTAATCTGCTAATCTGCTTCCCTGCTTCCAGACTCTGCCCACTCATCTCAGCTTT 2580  
Db CTGCTAATCTGCTAATCTGCTTCCCTGCTTCCAGACTCTGCCCACTCATCTCAGCTTT 6689  
QY 2581 CTTTCAAGAGTCAAGATGATGATGAGAGACCCCTCTGCTGCTTCTGTTCTTCCATG 2640  
Db CTTTCAAGAGTCAAGATGATGATGAGAGACCCCTCTGCTGCTTCTGTTCTTCCATG 6749  
QY 2641 ATTTCACCTGCACTCTGATTAAGTCCAGCTTGAACACAGCCTTACCAATCTTGCAGG 2700  
Db ATTTCACCTGCACTCTGATTAAGTCCAGCCTTGAACACAGCCTTACCAATCTTGCAGG 6809  
QY 2701 ATCTATGCTTATCTTCCATCTCTTATGCTATGCTTATGCTTATGCTTATGCTTATGCTT 2760  
Db ATCTATGCTTATCTTCCATCTCTTATGCTATGCTTATGCTTATGCTTATGCTTATGCTT 6869  
QY 2761 GCTATGCTGACCTCTGTTCTTCTTCCATTTTTTTTAAATTTTAAATTTGATATA 2820  
Db GCTATGCTGACCTCTGTTCTTCTTCCATTTTTTTTAAATTTTAAATTTGATATA 6929  
QY 2821 TTTATGGATTAAGTAAATCTTTTATGATGATGATGATGATGATGATGATGATGATGATG 2880  
Db TTTATGGATTAAGTAAATCTTTTATGATGATGATGATGATGATGATGATGATGATGATG 6989  
QY 2881 CTTTAAAGGATTAATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2940  
Db CTTTAAAGGATTAATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 7049  
QY 2941 TCCGCTGACTTCTTGCCCTGCGGATTTATCATCACTGGAATGATGATGATGATGATGATG 3000  
Db TCCGCTGACTTCTTGCCCTGCGGATTTATCATCACTGGAATGATGATGATGATGATGATG 7109

Page 4

QY	3001	AGTAATTTCTCAACCATCCGCTGA	CTTCTTGCCCCCTGGGTATTCA	TCACCTGATGATGT	3060
Db	7110	AGTAATTTCTCAACCATCCGCTGA	CTTCTTGCCCCCTGGGTATTCA	TCACCTGATGATGT	7169
QY	3061	GCATGTGACCCCTTAAGTAATTTCT	CACCATCGCGTGA	CTTCTTGCCCCCTCATCCTTCT	3120
Db	7170	GCATGTGACCCCTTAAGTAATTTCT	CACCATCGCGTGA	CTTCTTGCCCCCTCATCCTTCT	7229
QY	3121	GAGGCTTCATGTTCATTCATTCOCA	CACTCTCATGTGTGACATTTATTAGCTCT	3180	
Db	7230	GAGGCTTCATGTTCATTCATTCOCA	CACTCTCATGTGTGACATTTATTAGCTCT	7289	
QY	3181	ACTTATTAAGTAATGATGATCAAT	ATTGTCTTCTGTGCTCTGTGTTTACTATGAT	3240	
Db	7290	ACTTATTAAGTAATGATGATCAAT	ATTGTCTTCTGTGCTCTGTGTTTACTATGAT	7349	
QY	3241	AATGGCCCCAGTTCTATCTAGAGCT	GCTGCAAAAAGSCATGATTCTTTTATGGC	3300	
Db	7350	AATGGCCCCAGTTCTATCTAGAGCT	GCTGCAAAAAGSCATGATTCTTTTATGGC	7409	
QY	3301	TATGTTCTTCCCAATTTGATGAT	TAAGAACACCTGCGCTGCTTACTCTATTTGGAT	3360	
Db	7410	TATGTTCTTCCCAATTTGATGAT	TAAGAACACCTGCGCTGCTTACTCTATTTGGAT	7469	
QY	3361	ACTAATTCCTAGGCTTCTTGATGT	CTTCTCTCTCTCAACCATCAAAATCTCATTTTGA	3420	
Db	7470	ACTAATTCCTAGGCTTCTTGATGT	CTTCTCTCTCTCAACCATCAAAATCTCATTTTGA	7529	
QY	3421	TACACCTCTTCAAGAGGGCTTCTG	ACCACTTGCGTGAATTAGCCCTTCAACCAT	3480	
Db	7530	TACACCTCTTCAAGAGGGCTTCTG	ACCACTTGCGTGAATTAGCCCTTCAACCAT	7589	
QY	3481	GATTACTCTCTAGCACATCACTG	CCCATTTTATTCATGTACAGGTCAAAATCTGAA	3540	
Db	7590	GATTACTCTCTAGCACATCACTG	CCCATTTTATTCATGTACAGGTCAAAATCTGAA	7649	
QY	3541	CACCTGATTTGTTTATTTTCTGAC	CTCTCTCATCTAGAGAAATCTACTAGAGCGGAG	3600	
Db	7650	CACCTGATTTGTTTATTTTCTGAC	CTCTCTCTCATCTAGAGAAATCTACTAGAGCGGAG	7709	
QY	3601	ATTTTATCTGCTTGTATCAGGAT	AGCTTCAAAACGACCTTGATACAGAGTGTGTCA	3660	
Db	7710	ATTTTATCTGCTTGTATCAGGAT	AGCTTCAAAACGACCTTGATACAGAGTGTGTCA	7769	
QY	3661	AATGATTTTCTTAAACAAATGAA	CAATTAATAAGATGATCTTTTGAAGATGAAGCTTT	3720	
Db	7770	AATGATTTTCTTAAACAAATGAA	CAATTAATAAGATGATCTTTTGAAGATGAAGCTTT	7829	
QY	3721	CCACACTACAGAGTCATTCAAGAT	TGACAAATCATAGAAATTAACAGAAATTTGATGCTTTG	3780	
Db	7830	CCACACTACAGAGTCATTCAAGAT	TGACAAATCATAGAAATTAACAGAAATTTGATGCTTTG	7889	
QY	3781	TGCATATCAGAGAAAGAGGTGGA	AGTTGTCAAGATATCATGATGTACCGTCTCGCC	3840	
Db	7890	TGCATATCAGAGAAAGAGGTGGA	AGTTGTCAAGATATCATGATGTACCGTCTCGCC	7949	
QY	3841	TCCTCAAAACAATCTGCAAGTCC	ACAGTGAAGAAAGTAACTGATGAGAGCTT	3900	
Db	7950	TCCTCAAAACAATCTGCAAGTCC	ACAGTGAAGAAAGTAACTGATGAGAGCTT	8009	
QY	3901	TTTCAAAACATTTTAAAAAGTCT	TAAAACTCTTAAGAAAGCAAAATTTAATAGTCAAG	8069	
Db	8010	TTTCAAAACATTTTAAAAAGTCT	TAAAACTCTTAAGAAAGCAAAATTTAATAGTCAAG	8069	
QY	3961	AAGTGAATTAACATGAAATGCT	TGAACAGAGTAATGAGCTTAAGCAAAAGTTAGAGCAT	4020	
Db	8070	AAGTGAATTAACATGAAATGCT	TGAACAGAGTAATGAGCTTAAGCAAAAGTTAGAGCAT	8129	
QY	4021	GTTAGTATTAATGTCTTGAAGAG	AGAGCTCTCTGCTTTCAAGAGACAGAAATTTGGG	4080	
Db	8130	GTTAGTATTAATGTCTTGAAGAG	AGAGCTCTCTGCTTTCAAGAGACAGAAATTTGGG	8189	



```
Db      9270  ACCCTATCGCTTTCTAAATGAAAGCTGTTTGGTCAGATCTAGCAATTAATTCCTCTCT 9329
Qy      5221  CCACCTATAGCTTCCCTGTAACTCGGTAGGTAATTTGGTTATAGGCTATAGATG 5280
Db      9330  CCACCTATAGCTTCCCTGTAACTCGGTAGGTAATTTGGTTATAGGCTATAGATG 9389
Qy      5281  GAAACACCTGAATGATTTCTGTCCATGACGACATTTCACTTCATGATTTGTATGTAAG 5340
Db      9390  GAAACACCTGAATGATTTCTGTCCATGACGACATTTCACTTCATGATTTGTATGTAAG 9449
Qy      5341  ATACGATTTGCTAGGTTGTAGAAACCCATATGGGCTTAATATTTGTCATCAGTTT 5400
Db      9450  ATACGATTTGCTAGGTTGTAGAAACCCATATGGGCTTAATATTTGTCATCAGTTT 9509
Qy      5401  GAAAGCTGTATAGCAAGCAAACTACATATTTTCTGCTGCTCTCTCTCTCTCTCT 5460
Db      9510  GAAAGCTGTATAGCAAGCAAACTACATATTTTCTGCTGCTCTCTCTCTCTCTCT 9569
Qy      5461  CTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5520
Db      9570  CTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 9629
Qy      5521  CATAAACCCATTGTATATAGAAATGCTGTATATTTGCTCATCTCACTCCCTCTTGG 5580
Db      9630  CATAAACCCATTGTATATAGAAATGCTGTATATTTGCTCATCTCACTCCCTCTTGG 9689
Qy      5581  TCCTCGAGCTGCGGCTTTAGACTTTTATAGAGACGCGAGCATGTGAAGAAACTGTCA 5640
Db      9690  TCCTCGAGCTGCGGCTTTAGACTTTTATAGAGACGCGAGCATGTGAAGAAACTGTCA 9749
Qy      5641  GTGCTAGGCTGAATCTGTGTATCCAGATTTCTAGAAAAGTATCTAGTCAGAGTTG 5700
Db      9750  GTGCTAGGCTGAATCTGTGTATCCAGATTTCTAGAAAAGTATCTAGTCAGAGTTG 9809
Qy      5701  ATTACAGATATAGCAAAATCTATTTTCTAGGTAATTTCTGATGCTGCGGCTTATA 5760
Db      9810  ATTACAGATATAGCAAAATCTATTTTCTAGGTAATTTCTGATGCTGCGGCTTATA 9869
Qy      5761  ACTGCTGTCATCCAGCTATTTCTCTCCACCTTTCTTGTGCAATAACAAGGCAACT 5820
Db      9870  ACTGCTGTCATCCAGCTATTTCTCTCCACCTTTCTTGTGCAATAACAAGGCAACT 9929
Qy      5821  TCCGCAAAATCACTGCTGAGACGATGATCTGTCAGCTCCCTTTGGAAATCGTAGGA 5880
Db      9930  TCCGCAAAATCACTGCTGAGACGATGATCTGTCAGCTCCCTTTGGAAATCGTAGGA 9989
Qy      5881  TCAGATCTTTGACCATGTATATATGATGCTTTCAATCCAAAGAGAAAGCAATTGGGA 5940
Db      9990  TCAGATCTTTGACCATGTATATATGATGCTTTCAATCCAAAGAGAAAGCAATTGGGA 10049
Qy      5941  GTACACTCTTAAGTATCCAGAAATTCCTGCTGTAATTTCTTCCAGAGAAAGCAATT 6000
Db      10050  GTACACTCTTAAGTATCCAGAAATTCCTGCTGTAATTTCTTCCAGAGAAAGCAATT 10109
Qy      6001  CCTGTATATTTTTTTTACAGGCATATGAATAAATCTATATTTTGGACATTTGAC 6060
Db      10110  CCTGTATATTTTTTTTACAGGCATATGAATAAATCTATATTTTGGACATTTGAC 10169
Qy      6061  TTTTTCCTTTCTAGAAATTTCTAAACCTGTGCAATTTGGTGGACATTTGACATTT 6120
Db      10170  TTTTTCCTTTCTAGAAATTTCTAAACCTGTGCAATTTGGTGGACATTTGACATTT 10229
Qy      6121  TTTTCCCATATCCCTATCTTTTACAGAGATTTTCTCTGCTGTTCACTTAACTTGTAT 6180
Db      10230  TTTTCCCATATCCCTATCTTTTACAGAGATTTTCTCTGCTGTTCACTTAACTTGTAT 10289
Qy      6181  GGGTCAGTCTTTCTTCTCATCTCTTTTCAGGGGCTGGAAGGAGAGAGACAGAGA 6240
Db      10290  GGGTCAGTCTTTCTTCTCATCTCTTTTCAGGGGCTGGAAGGAGAGAGACAGAGA 10349
Qy      6241  GCTGTACTGACAGAGCGGTGCTGATTTGCTGACCGGTGTAAGCTGGCTATATAAGAG 6300
Db      10350  GCTGTACTGACAGAGCGGTGCTGATTTGCTGACCGGTGTAAGCTGGCTATATAAGAG 10409
Qy      6301  ACCCTACAGGCTTAGCAGAGAAACGCTCAGAGATTTCTGACATATCTTTACCGGAGAA 6360
Db      10410  ACCCTACAGGCTTAGCAGAGAAACGCTCAGAGATTTCTGACATATCTTTACCGGAGAA 10469
Qy      6361  GAGCAAAAGTACGCTCAAAGCCGAGACAGCTCCTCTGCGGCAATTTCTTCTCTCT 6420
Db      10470  GAGCAAAAGTACGCTCAAAGCCGAGACAGCTCCTCTGCGGCAATTTCTTCTCTCT 10529
Qy      6421  GCGAATTCGAAGCTGTAAATTAAGATGTGCAAGAGGCTTGCAAGCTCTGCGGCTTTCG 6480
Db      10530  GCGAATTCGAAGCTGTAAATTAAGATGTGCAAGAGGCTTGCAAGCTCTGCGGCTTTCG 10589
Qy      6481  TTGAGAGGTAAGTTCCTTTCAGCCATTAACCATATTAACCTTTTGGCTAGACTTCTC 6540
Db      10590  TTGAGAGGTAAGTTCCTTTCAGCCATTAACCATATTAACCTTTTGGCTAGACTTCTC 10649
Qy      6541  AGTTATTTACAGTTGTAATTAACCTTAAGTCTGAGCAATTAGAAACAGTGTGTCAG 6600
Db      10650  AGTTATTTACAGTTGTAATTAACCTTAAGTCTGAGCAATTAGAAACAGTGTGTCAG 10709
Qy      6601  GAGAGCAGACTTTCTAATCTTCTCCAGACTAGCTAGATATTTGACTTAAGACATGT 6660
Db      10710  GAGAGCAGACTTTCTAATCTTCTCCAGACTAGCTAGATATTTGACTTAAGACATGT 10769
Qy      6661  GCTCCCAAAATTTGAGCCCTTATGCTGTTTGTGTGAGACCTCAGTTTGAAGACTTTC 6720
Db      10770  GCTCCCAAAATTTGAGCCCTTATGCTGTTTGTGTGAGACCTCAGTTTGAAGACTTTC 10829
Qy      6721  TATCTTTAAGCCAGCTCTAAGAAAGCTAGTTTAAATTAAGAAAGCAGATGAGGTTTGA 6780
Db      10830  TATCTTTAAGCCAGCTCTAAGAAAGCTAGTTTAAATTAAGAAAGCAGATGAGGTTTGA 10889
Qy      6781  GCTATGTACAGTATCTGTATATATCTCCATCTGTATCTACTGCTATATTTAGACATCCC 6840
Db      10890  GCTATGTACAGTATCTGTATATATCTCCATCTGTATCTACTGCTATATTTAGACATCCC 10949
Qy      6841  TGAAGTACATTAAGAGCTGCTGAGGCTTCTGATTTGATGCTTAACAACCTGTTTACAGA 6900
Db      10950  TGAAGTACATTAAGAGCTGCTGAGGCTTCTGATTTGATGCTTAACAACCTGTTTACAGA 11009
Qy      6901  AAGGTACCCCAAGATGAGGTTTGGCTCATCATCAAGAAAGCACTATGCTTTCGTTGCG 6960
Db      11010  AAGGTACCCCAAGATGAGGTTTGGCTCATCATCAAGAAAGCACTATGCTTTCGTTGCG 11069
Qy      6961  TGGTGAGATTAATCTTCACTCTCTATGTTCTTATAGCAAAATGTTAACAATGAGATAGAT 7020
Db      11070  TGGTGAGATTAATCTTCACTCTCTATGTTCTTATAGCAAAATGTTAACAATGAGATAGAT 11129
Qy      7021  TTTAAGCAGATCTTCTATCTCTGCCCCATCTTAAGTCTTGAAGTGTCTCATAT 7080
Db      7081  GAGTTTGGTGAAGATTTGATCATTAACAATCAGTTAATAGTTTGTGAGAAATCTCA 7140
Qy      11190  GAGTTTGGTGAAGATTTGATCATTAACAATCAGTTAATAGTTTGTGAGAAATCTCA 11249
Qy      7141  TCTTAAGACATTTGTTTGTAAATATACCTCCCTGATTTTAAAAAGCCTTACAGACA 7200
Db      11250  TCTTAAGACATTTGTTTGTAAATATACCTCCCTGATTTTAAAAAGCCTTACAGACA 11309
Qy      7201  TACAGCTATTCATTTGTTTGTGTTGTTGTTCAAAAAAGTATTAAGAAATGCAATTCAGA 7260
Db      11310  TACAGCTATTCATTTGTTTGTGTTGTTGTTCAAAAAAGTATTAAGAAATGCAATTCAGA 11369
Qy      7261  AAGATCATATTTATGACAGTGAATAATTAACAATAATGAGCATATTAATTAATCTACTA 7320
Db      11370  AAGATCATATTTATGACAGTGAATAATTAACAATAATGAGCATATTAATTAATCTACTA 11429
Qy      7321  ATCTTGACGTCAAAAGTAAAGTCAACCTTAAGGTAATCACTGCTTTCTTATGCGAC 7380
Db      11430  ATCTTGACGTCAAAAGTAAAGTCAACCTTAAGGTAATCACTGCTTTCTTATGCGAC 11489
```

```
QY 7381 TGCAGATAGAAATACCAAAATTTATTTGGAAATATCTCAGAAAACATAATTTTT 7440
Db 11490 TGCAGATAGAAATACCAAAATTTATTTGGAAATATCTCAGAAAACATAATTTTT 11549
QY 7441 ATGTACTATTAACATTTACTTTCCAAATATCTGTCAATTCAGAGATGGAAGTACG 7500
Db 11550 ATGTACTATTAACATTTACTTTCCAAATATCTGTCAATTCAGAGATGGAAGTACG 11609
QY 7501 ATGGCTCTTTAAAGAGAGAGGGCTGCGAGAGATATCATTAATTAAGTTCTT 7560
Db 11610 ATGGCTCTTTAAAGAGAGAGGGCTGCGAGAGATATCATTAATTAAGTTCTT 11669
QY 7561 CTGACCTTCACGCTTAATTTCTGAATGAGTGAAGCAATTAATTCAGCTTCACTTAA 7620
Db 11670 CTGACCTTCACGCTTAATTTCTGAATGAGTGAAGCAATTAATTCAGCTTCACTTAA 11729
QY 7621 CTGCAATGAAATGAAACCGTACAAAATACAAAGTCTCAGAGAAAGTTATCTCTGG 7680
Db 11730 CTGCAATGAAATGAAACCGTACAAAATACAAAGTCTCAGAGAAAGTTATCTCTGG 11789
QY 7681 TAAATATTTTGCAGAAACAGATAAAAGATATACAGCTCTGCTCTCAAGATTAGC 7740
Db 11790 TAAATATTTTGCAGAAACAGATAAAAGATATACAGCTCTGCTCTCAAGATTAGC 11849
QY 7741 AGCTAATCTAGAGAGTAACTCTATGTGACAGAGATGAATCTGCTCTTCCCTTCTCT 7800
Db 11850 AGCTAATCTAGAGAGTAACTCTATGTGACAGAGATGAATCTGCTCTTCCCTTCTCT 11909
QY 7801 CAATTAATGCAAAATCATCTAGTCCAAACATCTTTACCAAGCTGAGGCTCCAGAG 7860
Db 11910 CAATTAATGCAAAATCATCTAGTCCAAACATCTTTACCAAGCTGAGGCTCCAGAG 11969
QY 7861 AGCCATTCGCTTCTCAAGGTCAATAGTGTGGTGAAGTTAGCAACCAATCTAGAAATTC 7920
Db 11970 AGCCATTCGCTTCTCAAGGTCAATAGTGTGGTGAAGTTAGCAACCAATCTAGAAATTC 12029
QY 7921 CTGACTCCAGTAATCTCTGAAGTCAATTTGTTTTTATTTTATGTTTTTATTAAGAA 7980
Db 12030 CTGACTCCAGTAATCTCTGAAGTCAATTTGTTTTTATTTTATGTTTTTATTAAGAA 12089
QY 7981 TACTTCTAGACACATTAACCCCTGCAATTGATTAATTAATCTAGAGTCTGAGTATC 8040
Db 12090 TACTTCTAGACACATTAACCCCTGCAATTGATTAATTAATCTAGAGTCTGAGTATC 12149
QY 8041 C 8041
Db 12150 C 12150
```

```
PR 29-MAR-2000; 2000US-0193480P.
PR 15-MAY-2000; 2000US-0205230P.
PR 09-JUN-2000; 2000US-0211315P.
PR 25-JUL-2000; 2000US-0220534P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Lillie J, Xu Y, Wang Y, Steinmann K,
XX WPI; 2001-451856/48.
XX
XX New peptide useful as a marker for the diagnosis of breast cancer.
XX
XX Claim 1; Page 3683-3684; 3695pp; English.
XX
XX The invention relates to human breast cancer expressed polynucleotides
XX (AA07544-AL026789) and methods of assessing whether a patient is
XX afflicted with breast cancer by examining the correlation between the
XX expression of certain markers and the cancerous state of breast cells.
XX The polynucleotides and encoded polypeptides are potential markers for
XX detecting, diagnosing, monitoring, characterizing treating and
XX potentially preventing breast cancer. The polynucleotides and encoded
XX polypeptides are also useful for isolating compounds with cytostatic
XX activity
XX
SQ Sequence 1238 BP; 386 A; 265 C; 307 G; 274 T; 0 U; 6 Other:
XX
XX Query Match 3.5%; Score 277.6; DB 4; Length 1238;
XX Best Local Similarity 98.6%; Pred. No. 28-54;
XX Matches 280; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
QY 6205 CTTTCAGGGGCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6264
Db 58 CTCTAGAGGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 117
QY 6265 GATTGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6324
Db 118 GATTGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1177
QY 6325 CGCTCAGAGAGATTCTGACATATTTTATCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 6384
Db 178 CGCTCAGAGAGATTCTGACATATTTTATCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 237
QY 6385 AGCCAGAGCTCTCTCCGCGATTTCTTCTGCTGCGAATTCAGAGCTGTTAAATAG 6444
Db 238 AGCCAGAGCTCTCTCTCCGCGATTTCTTCTGCTGCGAATTCAGAGCTGTTAAATAG 297
QY 6445 ATGTGCAAAAGGCTTGCAGAGTCTGCGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6488
Db 298 ATGTGCAAAAGGCTTGCAGAGTCTGCGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 341
```

```
RESULT 2
AAL26766
ID AAL26766 standard; cDNA; 1238 BP.
XX
XX AAL26766;
XX
XX 07-DEC-2001 (first entry)
XX
XX Human breast cancer expressed polynucleotide 19223.
XX
XX DE Human; breast cancer; cell marker; cytostatic; ss.
XX
XX KW Homo sapiens.
XX
XX OS MO200151628-A2.
XX
XX PN 19-JUL-2001.
XX
XX PD 10-JAN-2001; 2001WO-US000798.
XX
XX PF 14-JAN-2000; 2000US-0176077P.
XX
XX PR 14-MAR-2000; 2000US-0189167P.
XX
XX PD 24-MAR-2000; 2000US-0192099P.
```

```
RESULT 3
ABX63113
ID ABX63113 standard; cDNA; 3165 BP.
XX
XX ABX63113;
XX
XX 25-FEB-2003 (first entry)
XX
XX DE Human cDNA #113 differentially expressed in activated vascular tissue.
XX
XX KW Human; gene; ss; vascular tissue; cytostatic; atherosclerosis; cardiant;
XX hypotensive; antidiabetic; gynaecological; vasotrophic; cerebroprotective;
XX gene therapy; vascular disease; cancer; coronary; artery disease;
XX hypertension; diabetes; pre-eclampsia; restenosis;
XX ischaemia-reperfusion injury; stroke.
XX
XX OS Homo sapiens.
XX
XX PN US2002137081-A1.
XX
XX PD 26-SEP-2002.
```



XX	08-JAN-2002; 2002US-00044090.	PF
XX		XX
PR	28-JUL-2000; 2000US-0222469P.	PR
PR	08-JAN-2001; 2001US-0260483P.	PR
XX		XX
PA	(BAND/) BANDMAN O.	PA
PI	Bandman O;	PI
PI		PI
XX	WPI, 2003-110597/10.	XX
DR		DR
PT	Combination for diagnosing, staging, treating, or monitoring the	PT
PT	progression of treatment of a vascular disease, e.g. atherosclerosis,	PT
PT	comprises several CDNAs that are differentially expressed in activated	PT
PT	vascular tissue.	PT
XX		XX
PS	Claim 1; Page; 18pp; English.	PS
XX		XX
CC	This invention relates to a combination comprising several CDNAs that are	CC
CC	differentially expressed in activated vascular tissue. The invention also	CC
CC	discloses a high throughput method for detecting differentially expressed	CC
CC	CDNAs in a sample. The CDNAs of the invention may have	CC
CC	antiatherosclerotic, cytoprotective, cardiatic, hypotensive; antidiabetic;	CC
CC	synaescological, vasotropic and cerebroprotective activities and may be	CC
CC	used in gene therapy. The CDNAs of the invention may be used in a high-	CC
CC	throughput methods for detecting differential expression of one or more	CC
CC	CDNAs in a sample, or screening several molecules or compounds to	CC
CC	identify a molecule or compound that specifically binds a CDNA of the	CC
CC	invention. A protein encoded by the CDNA may be used to screen several	CC
CC	molecules or compounds to identify a ligand that specifically binds to	CC
CC	the protein, or to produce or purify an antibody to the protein that can	CC
CC	be used to detect a protein in a sample or purify a natural or	CC
CC	recombinant protein from a sample. The nucleotides may be useful for	CC
CC	diagnosing, staging, treating, or monitoring the progression of treatment	CC
CC	of a vascular disease, e.g. atherosclerosis, cancer, coronary artery	CC
CC	disease, hypertension, diabetes, pre-eclampsia, ischaemia-reperfusion	CC
CC	injury, restenosis, or stroke. The CDNAs can also be used for large-scale	CC
CC	genetic or gene expression analysis of several new nucleic acid	CC
CC	molecules. Antibodies to the proteins encoded by the CDNAs are useful for	CC
CC	diagnosing pre-pathologic disorders, and chronic or acute diseases	CC
CC	associated with abnormalities in the expression, amount or distribution	CC
CC	of the protein. The present sequence represents a CDNA of the invention	CC
CC	that is differentially expressed in activated vascular tissue. Note: The	CC
CC	sequence data for this patent did not form part of the specification, but	CC
CC	was obtained in electronic format directly from USPTO at	CC
CC	http://seqdata.uspto.gov/sequence.html?DocID=20020137081	CC
XX		XX
SQ	Sequence 3165 BP; 889 A; 636 C; 672 G; 967 T; 0 U; 1 Other;	SQ
	Query Match 3.4%; Score 274.4; DB 7; Length 3165;	
	Best Local Similarity 97.9%; Pred. No. 1.7e-55;	
	Matches 278; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
QY	6205 CTTTCAGGGGCTGAGAGGACAGAGGACACAGAGAGCTGTACTGTCAGAGCGGTGTCT	6264
DB	63 CGTAGAGGGCTCGAGAGGACAGAGGAGGAGACAGAGAGCTGTACTGTCAGAGCGGTGTCT	122
QY	6265 GATTGGCTGACGGTGTGTAAGTGGGCTTTAAAGAAGCCCTTACAGAGCTTAAGAGGA	6322
DB	123 GATTGGCTGACGGTGTGTAAGTGGGCTTTAAAGAAGCCCTTACAGAGCTTAAGAGGA	182
QY	6325 CGTTCAGAGGATTCGACAAATATCTTTACCGGAGAAAGGCAAGTACGCTCAAGCCGA	6384
DB	183 CGTTCAGAGGATTCGACAAATATCTTTACCGGAGAAAGGCAAGTACGCTCAAGCCGA	242
QY	6385 AGCCACAGCTCCTCTGCGGCAATTTCTTTCTGCTTGCGAATTCAGAGCTTTAAATAG	6444
DB	243 AGCCACAGCTCCTCTGCGGCAATTTCTTTCTGCTTGCGAATTCAGAGCTTTAAATAG	302
QY	6445 ATGTGCAAGGGCTTGACAGTGTGCGGCTCTTGCTTGAGAG 6488	
DB	303 ATGTGCAAGGGCTTGACAGTGTGCGGCTCTTGCTTGAGAG 346	

XX	RESULT 4
XX	ID ADD12660 standard; cDNA; 3167 BP.
XX	AC ADD12660;
XX	DT 01-JAN-2004 (first entry)
XX	DE Human cDNA expressed during adipocyte differentiation #6.
XX	KM ss; Gene; human; adipocyte differentiation; obesity; type II diabetes;
XX	KM lipodystrophy; hyperinsulinaemia.
XX	CS Homo sapiens.
XX	OS US2003113720-A1.
XX	PD 19-JUN-2003.
XX	PF 30-JUL-2001; 2001US-00918624.
XX	PR 28-JUL-2000; 2000US-0222470P.
XX	PA (SCHE/) SCHEBIE X M.
XX	PA (SORN/) SORNASSE T.
XX	PI Schebye XM, Sornasse T;
XX	DR WPI; 2003-810888/76.
XX	PT Novel isolated cDNAs expressed in adipocyte differentiation useful for
XX	PT treating subject with disorder such as obesity, type II diabetes,
XX	PT lipodystrophy or hyperinsulinemia.
XX	PS Claim 1; SEQ ID NO 6; 105bp; English.
XX	CC The invention relates to an isolated cDNA expressed during adipocyte
XX	CC differentiation. The cDNA is useful for treating a subject with a
XX	CC disorder such as obesity, type II diabetes, lipodystrophy or
XX	CC hyperinsulinaemia. The nucleic acid is useful for a high throughput
XX	CC method of using a cDNA to screen several molecules or compounds to
XX	CC identify a ligand which specifically binds the cDNA which involves
XX	CC combining the nucleic acid with several molecules or compounds under
XX	CC conditions to allow specific binding, and detecting specific binding
XX	CC between each cDNA and at least one molecule or compound, thus identifying
XX	CC a ligand that specifically binds to each cDNA. The several molecules or
XX	CC compounds are chosen from DNA molecules, RNA molecules, peptide nucleic
XX	CC acid molecules, mimetics, peptides, transcription factors, repressors and
XX	CC regulatory proteins. The present sequence represents a human cDNA
XX	CC expressed during adipocyte differentiation.
XX	SQ Sequence 3167 BP; 877 A; 625 C; 671 G; 970 T; 0 U; 24 Other;
QY	Query March 3 4% Score 274.4; DB 9; Length 3167;
D6	Best Local Similarity 97.9%; Pred. No. 1.7e-53;
Matches 278; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
QY	6205 CTTTCAGGGGCTCGAGAGGCAGAGGAGACAAGAGACTGTGTACTGCAAGACGGTGTCT 6266
D6	63 CGTAGAGGGGCTCGAGAGGCAGAGGAGACAAGAGACTGTGTACTGCAAGAGGATCCTCT 122
QY	6265 GATTGGCTTGAGCGGTCGAGTCAGTGGGCTATAAAGAAGACCCTACAGGCTTACAGGAAGA 6324
D6	123 GATTGGCTTGAGCGGTCGAGTCAGTGGGCTATAAAGAAGACCCTACAGGCTTACAGGAAGA 182
QY	6325 CGCTCAGAGGATTCAGACAATATCTTTACCGGAGAGAGGCAAAGTAGTCCTAAAGCCGA 6384
D6	183 CGCTCAGAGGATTCAGACAATATCTTTACCGGAGAGAGGCAAAGTAGTCCTAAAGCCGA 242
QY	6385 AGCGAAGCTCTCTCGGCCGCAATTTCTTCTGTTCCGAATCCAAAGTGTTAATAG 6444

Db	243	AGGACACAGCTTCCTGCGGCATTTCTTCTCGTTCGGAATTCACACTGTAAATAG	346
OY	6445	ATGTGCAAGAGGCTTCAGGTCGTCCGCGCTTCTTCCTTGAGAG	6488
Db	303	ATGTGCAAGAGGCTTCGAGCTCGCGGCTTCTTCCTTGAGAG	346
RESULT 5			
ID	AAK94664	standard; cDNA; 1902 BP.	
XX	AAK94664		
AC	AAK94664;		
DT	06-NOV-2001	(first entry)	
XX			
DE	Human full-length cDNA, SEQ ID NO: 3657.		
XX			
KW	Human; full length cDNA; cDNA synthesis; oligo-capping; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	EPI130094-A2.		
PD	05-SEP-2001.		
XX			
PF	07-JUL-2000; 2000EP-00114089.		
XX			
PR	08-JUL-1999; 99JP-00194486.		
PR	11-JAN-2000; 2000JP-00118774.		
ER	02-MAY-2000; 2000JP-00183765.		
XX			
PA	(HELI-) HELIX RES INST.		
XX			
PI	Ota T, Nishikawa T, Isogai T, Hayashi K, Ichii S, Kawai Y,		
PI	Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;		
DR	WPI; 2001-524255/56.		
Claim 8; SEQ ID NO 3657; 1380bp + Sequence Listing; English.			
830 Primers useful for synthesizing full length cDNA clones and their use			
in genetic manipulation.			
The invention relates to primers for synthesizing full length cDNA			
clones. 830 cDNA molecules encoding a human protein have been isolated			
and nucleotide sequences of 5' and 3' ends of the cDNA molecules have			
been determined. Primers for synthesizing the full length cDNA are useful			
for clarifying the function of the protein encoded by the cDNA. The full			
length clones were obtained by construction of full length enriched cDNA			
libraries that were synthesized by the oligo-capping method. The primers			
enable the production of the full length cDNA easily without any special			
methods. The present sequence is a full length human cDNA of the			
invention. Note: The sequence data for this patent did not form part of			
the printed specification, but was obtained in CD-ROM format directly			
from EPO			
Sequence 1902 BP; 576 A; 333 C; 440 G; 553 T; 0 U; 0 Other;			

QY	1422	AAAGAGCTCAGGGGGAAACACAGCAAGGCGCTGTCTATTGCAATTTCTTTATCTCTCTC	1484
Db	386	AATAGACTAGAGTGTGGCAAAACCGA-GCCAGGCTGTCCAGACATTCTTGGTCTCTC--	442
QY	1485	TCTCTATGTATAGCAATCTTTTCCCTCTGAGTATGGGGCAGACTCTTCTTCAATGAGGT	1544
Db	443	-----TGTGTACATTTCTCTCCCTCCGTGTATAGGGCACACCCCACTGGAAATGAGGAT	496
QY	1545	CTTCAAGGAGAAAGGAGAAAGTGG-----CCTTTTAATTTATGGCTTGCTTGG	1596
Db	497	CTTCAAGGAGAAAGGAGAAAGGAGAGTAACTTTCAGAAAGTTTATGACTGTGCTGTG	556
QY	1597	GGGA-AGAGSAGTTTCTAATTTCTATGACCATCTTGGGAAAGAAATTCGTGTTTCTGT	1655
Db	557	GGAAAGAGGAGTTCAATGTCTATATGACCTGCGCTTGGAAACAGAAATTTATTTCTAT	616
QY	1656	GA-CTTGCTTTCATGAAAGAAAGAGAGCTAAAGGCGAGAGGCGAGAGATGCTAGAAAG	1714
Db	617	GAATCTACATATGGGGAGAAAGGGGAGCAACAAACACAGGGGCTGGAGAAAGGCGACAG	676
QY	1715	AGACTGTGGTGCTTGTAGGGGCTTCGCGCTCTCTTAACTTCAAGTACTCTTATGCAATAC	1774
Db	677	AGGCTTTGC---TTTGGGGCTTTCTAATTTCTTTAGTTCAAAG---TCCITGGAGTGC	729
QY	1775	CAAGCAGCTATACCTTGGCATATGGTTTTTGTAGACTCTTAACACTGCAATCATATGTAAC	1834
Db	730	CAAGGATTAATTTTGGAGATTCCTCTTCAAGCACAACAACTAAAAAAATGAAAT	789
QY	1835	CCTCT 1839	
Db	790	CATTT 794	

RESULT	6	
ABT000010	10	
Continuation (11 of 15) of ABT00010 from base 1000001 (Human neurogranin 1 gene.)		
WP	Sequence split into 15 fragments	LOCUS ABT00010 Accession ABT00010
WP	Fragment Name	Begin
WP	ABT00010_00	1
WP	ABT00010_01	100001
WP	ABT00010_02	200001
WP	ABT00010_03	300001
WP	ABT00010_04	400001
WP	ABT00010_05	500001
WP	ABT00010_06	600001
WP	ABT00010_07	700001
WP	ABT00010_08	800001
WP	ABT00010_09	900001
WP	ABT00010_10	1000001
WP	ABT00010_11	1100001
WP	ABT00010_12	1200001
WP	ABT00010_13	1300001
WP	ABT00010_14	1400001
WP	ABT00010_15	1503841

Query Match	2.1k	Score 170.2	DB 6	Length 110000
Best Local Similarity	68.8%	Prd. No. 1.9e-28		
Matches 222	Conservative 1	Mismatches 104	Indels 0	Gaps 0
QY	3006	TTTCTCACACCCGCTACTTCTTGCCCTCGATATTCATACCTGAATGATGTCATT	3065	
Db	13790	TTTGTATACATGATATATACCTCGTCTTTCAGTGTAAACATCACCAATAGTGATACCT	13845	
QY	3066	GTAACCCCTTAGTATATTTCTCACATCCGCTACTTCTTGCCCCCTATCCTTGAGGC	3125	
Db	13850	GTAACCTATATAGTACTTGATCTCTCTCAACCTCCCTCCACCTCCCATCTTTTGTAGTC	13905	
QY	3126	TCCATTGTCCATCAATTCACACCTTCACTATATGTGTACATATTTAGTCCCTACTTA	3185	
Db	13910	TCCATGTCTATATATTCACCCCTATATGTGCATAGTGACTATATTTAGTCTTACTTA	13965	
QY	3186	TAAAGTATACATCAATATTTGTCTTCTGTCTGTCTGTCTGTTTACTTATGTAAATGG	3245	

[illegible]

RESULT 7  
ABT01503\_10  
Continuation (11 of 15) of ABT01503 from base 1000001 (Human neuregulin 1 gene.)  
MP Sequence split into 15 fragments  
LOCUS ABT01503 Accession Abt01503  
Frag 1

Query Match	2.1%;	Score 170.2;	DB 6;	Length 110000;
Best Local Similarity	68.8%;	Pred. No. 1.9e-28;		
Matches 232; Conservative	1;	Mismatches 104;	Indels 0;	Gaps 0;

WP	AAKS5240_07	700001	810000
WP	AAKS5240_08	800001	910000
WP	AAKS5240_09	900001	1010000
WP	AAKS5240_10	1000001	1110000
WP	AAKS5240_11	1100001	1210000
WP	AAKS5240_12	1200001	1310000
WP	AAKS5240_13	1300001	1400000
WP	AAKS5240_14	1400001	1503900

Query Match	2.1%;	Score 169.6;	DB 4;	Length 110000;
Best Local Similarity	68.8%;	Pred. No. 2.6e-28;		
Matches 232, Conservative	0;	Mismatches 105;	Indels 0;	Gaps 0;

Db	13790	TTTGTTCACAGCATATACCTCGTCTTTCATGATGAACCATCATCAATATAGTACTT	13849
		RESULT 9	
		AAK96733_10	
Qy	3066	GTACCCCTTAGTAATTTCTACCAACGCGTACTTTGGCCCCCTATCCTTCTGAGGC	3125
		Continuation (11 of 15) of AAK96733 from base 1000001 (Human neurogranin-1 gene.)	
		WP Sequence split into 15 fragments	
Db	13850	GTACCTATTACGTAAGTTCATCTCTCCACCCCTCCACCCCTCCACTTTTGTGATC	13905
		WP Fragment Name	
		AAK96733_00	Begin 1
Qy	3126	TCCATTGTCCATCATTCACACACTCTCATCTATGATGATACATTAATTAGCTCCTACTTA	3185
		AAK96733_01	100001
		AAK96733_02	210000
Db	13910	TCCAAATGTATATATTCACCCCTCTATGTCCAAATGTAATCATTAATTAGTCTTACTTA	13969
		AAK96733_03	200001
		AAK96733_04	300001
Qy	3186	TAAATGATTAACAGCAATATTTGTCTTTGTGTGTCTGTCTGTTTACTTATGATAATGG	3245
		AAK96733_05	400001
Db	13970	TAAATGAACACAGCAATATTATTAATCTTCGTTCTGACTTTTTCACCTTAAGATTAAG	14029
		AAK96733_06	500001
Qy	3246	CCCCAGTTCATCTAGCGTGTGCAAAAGCATGATTCATTCCTTTTAAATGGCTATGT	3305
		AAK96733_07	600001
		AAK96733_08	700001
Db	14030	CCTTCAGGTCATTCAGTGTGTGCAAAAGCATGATTCCTTTGTATGGCTGAGT	14089
		AAK96733_09	800001
Qy	3306	TCTTTCCCAATTAGATTAAGAAACATCCGCACTTGCT	3342
		AAK96733_10	900001
Db	14090	AGTATTCATGCTATATACGTTACATTTTCTTGAT	14126
		AAK96733_11	1000001
		AAK96733_12	1100000
		AAK96733_13	1200001
		AAK96733_14	1300001
		AAK96733_15	1400001
		AAK96733_16	1503900
		Query Match	
		2.1%;	Score 169.6; DB 4; Length 110000;

[illegible]

Fragment Name	Begin	End	LOCUS AAK95240	Accession AAK95240
AAK95240_00	1	110000		
AAK95240_01	100001	210000		
AAK95240_02	200001	310000		
AAK95240_03	300001	410000		
AAK95240_04	400001	510000		
AAK95240_05	500001	610000		
AAK95240_06	600001	710000		

Db 13969 TCCATGTCATATATCCACCCTGATGTCATGCTACTCATTTATTAGTCTTACTTA 14028  
Qy 3186 TAAATGATTAACATGCAATATTGCTCTTCTGCTGCTGCTTGTATTACTTAATGATGATG 3245  
Db 14029 TAAATGAACATGCAATATTGCTCTTCTGCTGCTGCTTGTATTACTTAATGATGATG 14088  
Qy 3246 CCCCCAGTCTATCTAGCTGCTGCAAAAGCAGATTCATCTCTTTTATGCTATGT 3305  
Db 14089 CCTCAGCTCCATCCATGCTGCTGCAAAAGCAGATTCATCTCTTTTATGCTATGT 14148  
Qy 3306 TCTTTCCCAATTAGATTAAGAAACTCGCACTTGT 3342  
Db 14149 AGTATTCATGAGTGTATATAGCTTACATTTTCTTGAT 14185

RESULT 10  
AAK86211  
ID AAK86211 standard; DNA; 7457 BP.  
XX AAK86211;  
AC  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:41023.  
XX  
KM Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KM cytosolic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
PN W0200157182-A2.  
PD  
XX 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001354.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180629P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189674P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-020515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0217496P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 16-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0232082P.  
PR 14-SEP-2000; 2000US-0233397P.  
PR 14-SEP-2000; 2000US-0233398P.  
PR 14-SEP-2000; 2000US-0233399P.  
PR 14-SEP-2000; 2000US-0233400P.  
PR 14-SEP-2000; 2000US-0233401P.  
PR 14-SEP-2000; 2000US-0233402P.  
PR 14-SEP-2000; 2000US-0233403P.  
PR 14-SEP-2000; 2000US-0233404P.  
PR 14-SEP-2000; 2000US-0233405P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 26-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0234999P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241121P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246529P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.



Sequence 401 BP; 98 A; 79 C; 62 G; 161 T; 0 U; 1 Other;

Query Match 2.1%; Score 167.8; DB 4; Length 401;  
Best Local Similarity 71.2%; Pred. No. 6.2e-29;  
Matches 220; Conservative 1; Mismatches 88; Indels 0; Gaps 0;

QY 3006 TTCTCAGCATCCGCTGCTCTTGGCCCGGGATTATCATCAGCATGATGTCATT 3065  
DB 88 TTGTTACATGATGATTAATCTCTGCTTTTCAGTACATCATCAATAGTACCTT 147  
QY 3066 GTACCCCTTAAGTAATTTCTCACCACCGCTGACTTTGCCCCCATCTCTGAGGC 3125  
DB 148 GTACCTATTAGTAAGTGTGCTGCTGACCCCTCCACCCCTCCATCTTTGAGTGC 207  
QY 3126 TCCATTGTCATGATTCACACACTCTACATCTATGTAACATTTTATAGCCCTACTTA 3185  
DB 208 TCCAAATGCTATTTATTCACCCCTCTATGTCATGTAATCTATTTAGTTCTTACTTA 267  
QY 3186 TAAGTATTAACATGATCAATTTGCTTTCTGCTGCTGCTTTTACTTATGATATAG 3245  
DB 268 TAAGTATTAACATGATCAATTTGCTTTCTGCTGCTGCTTTTACTTATGATATAG 3245  
QY 3246 CCCCAGTTCTATCTGAGGCTGCAAAAGCATGATTTCACTTTTATGCTATG 3305  
DB 328 CTTTCAAGTTCATCTCATCTGCTGCTGCAAAAGCATGATTTGATGAGCTGAGT 387  
QY 3306 TCTTTCCCA 3314  
DB 388 AGTATTCCA 396

RESULT 12

AAK97635 standard; DNA; 401 BP.

AAK97635;

17-DEC-2001 (first entry)

Human neuregulin gene single nucleotide polymorphism SNP8NRG1013903.

Human, neuregulin 1 gene; schizophrenia; gene therapy; SNP;

single nucleotide polymorphism; ds.

Homo sapiens.

MO200164877-A2.

07-SEP-2001.

28-FEB-2001; 2001MO-US006377.

28-FEB-2000; 2000US-00515716.

(DECO-) DECODE GENETICS EHF.

Stefansson H, Steinthorsdottir V, Gulcher JR;

WPI, 2001-514841/56.

Neuregulin 1 nucleic acids and proteins useful for diagnosing preventing

and treating schizophrenia.

Disclosure; Page 244; 756bp; English.

This sequence represents a single nucleotide polymorphism (SNP) from the human neuregulin 1 gene of the invention. The invention also relates to fragments or variants of the neuregulin 1 gene. The gene and its proteins may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate neuregulin 1 expression, such as schizophrenia. For example they may be used to treat disorders associated with decreased neuregulin 1 expression by rectifying mutations or deletions in a patient's genome that affect the activity of neuregulin 1

by expressing inactive proteins or to supplement the patients own

production of polypeptides. Additionally, the gene may be used to produce

the neuregulin 1 protein, by inserting the nucleic acids into a host cell

and culturing the cell to express the protein. The gene and its

complementary sequences may also be used as DNA probes in diagnostic

assays to detect and quantify the presence of similar nucleic acids in

samples, and therefore which patients may be in need of restorative

therapy. The protein may also be used as antigens in the production of

antibodies against neuregulin 1 and in assays to identify modulators of

neuregulin 1 expression and activity. The antibodies and antagonists may

also be used to down regulate expression and activity. The antibodies may

also be used as diagnostic agents for detecting the presence of

neuregulin 1 in samples

Sequence 401 BP; 98 A; 79 C; 62 G; 161 T; 0 U; 1 Other;

Query Match 2.1%; Score 167.8; DB 4; Length 401;

Best Local Similarity 71.2%; Pred. No. 6.2e-29;

Matches 220; Conservative 1; Mismatches 88; Indels 0; Gaps 0;

QY 3006 TTCTCAGCATCCGCTGCTCTTGGCCCGGGATTATCATCAGCATGATGTCATT 3065

DB 88 TTGTTACATGATGATTAATCTCTGCTTTTCAGTACATCATCAATAGTACCTT 147

QY 3066 GTACCCCTTAAGTAATTTCTCACCACCGCTGACTTTGCCCCCATCTCTGAGGC 3125

DB 148 GTACCTATTAGTAAGTGTGCTGCTGACCCCTCCACCCCTCCATCTTTGAGTGC 207

QY 3126 TCCATTGTCATGATTCACACACTCTACATCTATGTAACATTTTATAGCCCTACTTA 3185

DB 208 TCCAAATGCTATTTATTCACCCCTCTATGTCATGTAATCTATTTAGTTCTTACTTA 267

QY 3186 TAAGTATTAACATGATCAATTTGCTTTCTGCTGCTGCTTTTACTTATGATATAG 3245

DB 268 TAAGTATTAACATGATCAATTTGCTTTCTGCTGCTGCTTTTACTTATGATATAG 3245

QY 3246 CCCCAGTTCTATCTGAGGCTGCAAAAGCATGATTTCACTTTTATGCTATG 3305

DB 328 CTTTCAAGTTCATCTCATCTGCTGCTGCAAAAGCATGATTTGATGAGCTGAGT 387

QY 3306 TCTTTCCCA 3314

DB 388 AGTATTCCA 396

RESULT 13

ABT00912 standard; DNA; 401 BP.

ABT00912;

07-NOV-2002 (first entry)

Human neuregulin-1-associated gene 1 fragment SEQ ID NO: 941.

Human, neuregulin 1; neuregulin-1-associated gene 1; NRG1; NRG1A1;

schizophrenia; chromosome 8p12; single nucleotide polymorphism; SNP;

neuroleptic; gene therapy; splice variant; gene; ds.

Homo sapiens.

US2002045577-A1.

18-APR-2002.

28-FEB-2001; 2001US-00795668.

28-FEB-2000; 2000US-00515716.

(DECO-) DECODE GENETICS EHF.

Stefansson H, Steinthorsdottir V, Gulcher JR;



Db 731 CAAGCATATATTTGGAGTATCTCTTCAAGCCACACCA 772  
 Search completed: September 11, 2004, 11:15:15  
 Job time : 1843 secs

08-JUL-1999; 99JP-00194486.  
 11-JAN-2000; 2000JP-00118774.  
 02-MAY-2000; 2000JP-00183765.  
 (HELI-) HELIX RES INST.  
 Oka T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,  
 Makamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H,  
 WPI; 2001-524255/58.  
 830 Primers useful for synthesizing full length cDNA clones and their use  
 in genetic manipulation.  
 Claim 2; SEQ ID NO 147; 1380bp + Sequence listing; English.  
 The invention relates to primers for synthesizing full length cDNA  
 clones. 830 cDNA molecules encoding a human protein have been isolated  
 and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have  
 been determined. Primers for synthesizing the full length cDNA are useful  
 for clarifying the function of the protein encoded by the cDNA. The full  
 length clones were obtained by construction of full length enriched cDNA  
 libraries that were synthesised by the oligo-capping method. The primers  
 enable the production of the full length cDNA easily without any special  
 methods. The present sequence is the nucleotide sequence of the 5'-end of  
 a cDNA provided in the invention. Note: The sequence data for this patent  
 did not form part of the printed specification, but was obtained in CD-  
 ROM format directly from EPO  
 Sequence 847 BP; 272 A; 143 C; 199 G; 229 T; 0 U; 4 Other;

Query Match 2.1%; Score 166.2; DB 4; Length 847;  
 Best Local Similarity 66.3%; Pred No 2e-28;  
 Matches 386; Conservative 0; Mismatches 164; Indels 32; Gaps 9;

QY 1245 CAGGTTAATGAGCAAAACCTACCAATTATTTAATCAAGTTTATACATGACATGGGA 1304  
 Db 213 CAAGTTAATGAGCAAAAGGTATACAAA-----TTAATCAAGATTATGTGACCAAGA 267  
 QY 1305 GTCTTGAGAAATGAACCAACCAAGACCCAGGGGAAACGTCTGTTTTCGTGAGGTT 1364  
 Db 268 GCCTTCAAAATGAAAGACCCAAAGATGAGGAAAGCTATCTGTTTCT--GCTAAGTT 325  
 QY 1365 CGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1424  
 Db 326 TAACGTAGATGAGACCGCTGCTAGAAATGCGACTGACAAAAGGATATGATTGCTGAC 385  
 QY 1425 AAAGGACTCAGGGGGGAAACACAGCAAGGCTGTCTATTCAGATTCTTGTGATCTCTC 1484  
 Db 386 AATGACTGAGAGTGTGGCAACCGA-GCCAGGCTGTCCAGACACTTCTGTCTCTC-- 442  
 QY 1485 TCTCTATGATAGCATTTCTCTCTCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGG 1544  
 Db 443 -----TGTGTAACATTTCTCTCTCTGATGATGAGGAGGAGGAGGAGGAGGAGG 496  
 QY 1545 CTTCAAGGAGAAAGGAGAAAGTG-----CCTTTTATGATTATGCTTGTGCTTGG 1596  
 Db 497 CTTCAAGGAGAAAGGAGAAAGGAGAAAGGAGAAAGGAGAAAGGAGAAAGGAGAAAG 556  
 QY 1597 GGGA-AGAGAGGTTCTAGTTTCTATGACCATCTTGGGGAGAGAGAAATCTGGTTCTGT 1655  
 Db 557 GGAAGAGAGGATTCAGTGTCTATGACCTGCTTGGAGAGACGAATTTATTTCTCTAT 616  
 QY 1656 G-ACCTGCTTCTATGAGAAAGAGAGTAAAGGAGAGGAGAGGAGAGGAGAGGAGAG 1714  
 Db 617 GAACCTACTATGAGGAGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 673  
 QY 1715 AGACTTGCTGCTCTCTGAGGGGCTTCCGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 1774  
 Db 674 ACGAGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 730  
 QY 1775 CAAGCATATATCTTGGCATATGTTTCTGAGCTCAACA 1816





```
US-09-536-059-1
Sequence 1, Application US/09536059
Patent No. 6544737
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Chumakov, Ilya
APPLICANT: Bouguerelet, Lydie
APPLICANT: Cohen-Akenzie, Annick
TITLE OF INVENTION: GENOMIC SEQUENCE OF THE PURH GENE AND PURH-RELATED BIALLELIC
MARKERS.
FILE REFERENCE: GENSET.058AUS
CURRENT APPLICATION NUMBER: US/09/536,059
CURRENT FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: US 60/125,961
PRIOR FILING DATE: 1998-03-24
NUMBER OF SEQ. ID NOS: 24
SOFTWARE: Patent.pm
SEQ ID NO 1
LENGTH: 41684
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..2000
OTHER INFORMATION: 5'regulatory region
NAME/KEY: exon
LOCATION: 2001..2096
OTHER INFORMATION: exon 1
NAME/KEY: exon
LOCATION: 2433..2559
OTHER INFORMATION: exon 2
NAME/KEY: exon
LOCATION: 8092..8168
OTHER INFORMATION: exon 3
NAME/KEY: exon
LOCATION: 9600..9666
OTHER INFORMATION: exon 4
NAME/KEY: exon
LOCATION: 15178..15266
OTHER INFORMATION: exon 5
NAME/KEY: exon
LOCATION: 15924..16075
OTHER INFORMATION: exon 6
NAME/KEY: exon
LOCATION: 16759..16915
OTHER INFORMATION: exon 7
NAME/KEY: exon
LOCATION: 22309..22434
OTHER INFORMATION: exon 8
NAME/KEY: exon
LOCATION: 23277..23384
OTHER INFORMATION: exon 9
NAME/KEY: exon
LOCATION: 24841..24926
OTHER INFORMATION: exon 10
NAME/KEY: exon
LOCATION: 25957..26046
OTHER INFORMATION: exon 11
NAME/KEY: exon
LOCATION: 28700..28828
OTHER INFORMATION: exon 12
NAME/KEY: exon
LOCATION: 34699..34791
OTHER INFORMATION: exon 13
NAME/KEY: exon
LOCATION: 36679..36861
OTHER INFORMATION: exon 14
NAME/KEY: exon
LOCATION: 39014..39169
OTHER INFORMATION: exon 15
NAME/KEY: exon
LOCATION: 39456..39684
OTHER INFORMATION: exon 16

NAME/KEY: misc.feature
LOCATION: 39685..41684
OTHER INFORMATION: 3'regulatory region
NAME/KEY: allele
LOCATION: 6491
OTHER INFORMATION: 99-32284-107 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 15234
OTHER INFORMATION: 99-5602-372 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 15868
OTHER INFORMATION: 5-290-32 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 16729
OTHER INFORMATION: 99-22573-321 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 18311
OTHER INFORMATION: 99-22586-300 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 18572
OTHER INFORMATION: 99-22586-39 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 22906
OTHER INFORMATION: 99-5596-197 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 23175
OTHER INFORMATION: 5-293-76 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 23253
OTHER INFORMATION: 5-293-155 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 26106
OTHER INFORMATION: 5-294-285 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 30464
OTHER INFORMATION: 99-23454-317 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 30669
OTHER INFORMATION: 99-23454-105 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 31250
OTHER INFORMATION: 99-15528-333 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 35148
OTHER INFORMATION: 99-15798-86 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 36801
OTHER INFORMATION: 5-297-209 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 37286
OTHER INFORMATION: 99-32281-276 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 37536
OTHER INFORMATION: 99-32281-26 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 39321
OTHER INFORMATION: 5-298-376 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 39689
OTHER INFORMATION: 99-23460-169 : polymorphic base G or T
NAME/KEY: primer_bind
LOCATION: 6137..6157
OTHER INFORMATION: 99-32284.rp
NAME/KEY: primer_bind
LOCATION: 6577..6597
OTHER INFORMATION: 99-32284.pu complement
NAME/KEY: primer_bind
LOCATION: 14864..14882
OTHER INFORMATION: 99-5602.pu
NAME/KEY: primer_bind
LOCATION: 15292..15312
OTHER INFORMATION: 99-5602.rp complement
NAME/KEY: primer_bind
```

LOCATION: 15837..15855  
OTHER INFORMATION: 5-290.pu  
NAME/KEY: primer\_bind  
LOCATION: 16249..16266  
OTHER INFORMATION: 5-290.rp complement  
NAME/KEY: primer\_bind  
LOCATION: 16599..16617  
OTHER INFORMATION: 99-22573.rp  
NAME/KEY: primer\_bind  
LOCATION: 17030..17049  
OTHER INFORMATION: 99-22573.pu complement  
NAME/KEY: primer\_bind  
LOCATION: 18131..18150  
OTHER INFORMATION: 99-22586.rp  
NAME/KEY: primer\_bind  
LOCATION: 18592..18610  
OTHER INFORMATION: 99-22586.pu complement  
NAME/KEY: primer\_bind  
LOCATION: 22710..22727  
OTHER INFORMATION: 99-5596.pu  
NAME/KEY: primer\_bind  
LOCATION: 23100..23118  
OTHER INFORMATION: 5-293.pu  
NAME/KEY: primer\_bind  
LOCATION: 23130..23149  
OTHER INFORMATION: 99-5596.rp complement  
NAME/KEY: primer\_bind  
LOCATION: 23512..23530  
OTHER INFORMATION: 5-293.rp complement  
NAME/KEY: primer\_bind  
LOCATION: 25822..25840  
OTHER INFORMATION: 5-294.pu  
NAME/KEY: primer\_bind  
LOCATION: 26222..26241  
OTHER INFORMATION: 5-294.rp complement  
NAME/KEY: primer\_bind  
LOCATION: 30332..30352  
OTHER INFORMATION: 99-23454.rp  
NAME/KEY: primer\_bind  
LOCATION: 30754..30773  
OTHER INFORMATION: 99-23454.pu complement  
NAME/KEY: primer\_bind  
LOCATION: 30918..30935  
OTHER INFORMATION: 99-15528.pu  
NAME/KEY: primer\_bind  
LOCATION: 31390..31408  
OTHER INFORMATION: 99-15528.rp complement  
NAME/KEY: primer\_bind  
LOCATION: 34780..34799  
OTHER INFORMATION: 99-15798.rp  
NAME/KEY: primer\_bind  
LOCATION: 35215..35233  
OTHER INFORMATION: 99-15798.pu complement  
NAME/KEY: primer\_bind  
LOCATION: 36593..36610

Query Match 1.8%; Score 141; DB 4; Length 41684;  
Best Local Similarity 63.7%; Pred. No. 2,1e-27;  
Matches 375; Conservative 0; Mismatches 175; Indels 39; Gaps 9;  
QY 1162 CAAAGCTGAGTAGGAGTGAACCTGATACATTTCTTCCATCATAGGGTGGCA 1221  
DB 4727 CCAAGATTTTGAGGGGTGGAAGGTGATCC--TTTGCTCTCGATCGTAAGCAGC 4784  
QY 1222 ACCAAACTCCTATAGTAAAGACAGGT--TAAATAGAGCAAAACCTAACAAATTATTT 1279  
DB 4785 GCCAATATTTTATATAGAGAGACAGGTTTAAATAGAGAAAACATTAACAAATTTATTT 4844  
QY 1280 AATCAAGTTTACATGATGCGAGCTTTCAGAAATGAAGCCCAAGCCCAAGGCGGA 1339  
DB 4845 AA-CAAAGTTTACATGATGAGAGCTTCAG--AATGAGACCCCAAGACAGAGGAAA 4902  
QY 1340 ACTGCTGTTTTTTTGTGCTGAGGTTCATGAAGAATGATAGCATGTAGCATGTAGATT 1399

DB 4903 AACATC-CATTTATGTTTAGGTTCAACAAAGAAAGACAGAAAGTGAATGATT 4961  
QY 1400 AGACAAAAGATATGATCTAGTGTAAAGACTCAG--GGGGAAGACAGCAAGGCTGT 1457  
DB 4962 GACAGCAAGATATGAGTATGCTAGTAGACTGAGGTGAGAAAGCAGAAAGCCTGT 5021  
QY 1458 CTATTGATTTCTTCTGATCTCTCTCTATGT-----ATAGATTTCTTCTCT 1508  
DB 5022 CTGTCCAGATTTCTTCTGCTCTCTTAAATTTCTTCTCCACCCCTTAAGATCTCTGA 5081  
QY 1509 CCTGAGTATGGGCGACACTCTTCTTCATAGAGGTTCTTCAGAGGAGAAAG-- 1566  
DB 5082 CCTACTATATGGGCAAGAGATAGAGAGATTTCTTTAGCTCCAGCTCCTAGACAGAAAGCC 5141  
QY 1567 -----TGCCCTTTTATGATTTTATGCTTCTTCCGGGAAAGAGAGTTCT 1611  
DB 5142 AGTGAAAGTTAGATCATTAAGTTTAAATCTTATGATGAGCTTTGGGAAAAGAGTTT 5201  
QY 1612 AGTTTCATGACCCATCTTGGGGAAGAGAAATCTGTTCTGTGACTTCTTATGAA 1671  
DB 5202 AGTTTCTTAAGTCCCTGCGGGAAGAAATTTCTATTTCTGTGACTTCTAGG----GAA 5257  
QY 1672 GAAAGAGAGTTAAGAGCAGAGAGGCGAGATGTCAGAAAGACTT 1720  
DB 5258 GAATGAAGGCTGAGAGGCAAGAGGCGAGAGAAAGCTATATATATTT 5306

RESULT 3  
US-09-016-434-1091  
Sequence 1091, Application US/09016434  
Patent No. 6500938  
GENERAL INFORMATION:  
APPLICANT: Janice Au-Young  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
NUMBER OF SEQUENCES: 1490  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: word perfect 6.1 for windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/016,434  
FILING DATE: HERewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0002 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1091:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 800 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK

CLONE: 91216372  
US-09-016-434-1091

Query Match 1.6%; Score 125; DB 4; Length 800;  
Best Local Similarity 100.0%; Pred. No. 3.7e-24;  
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6364 GCAAGTAGCCTCAAGCCGGAAGCCACAGAGCTCCTCGCGGATTTCTTCTGCTGCG 6423  
DB 17 GCAAGTAGCCTCAAGCCGGAAGCCACAGAGCTCCTCGCGGATTTCTTCTGCTGCG 76

QY 6424 AATTCCAGCTGTTAATAATAGATGTGCAGAGGGCTTGAGAGTCTGCGGCTTTCTTGCTTG 6483  
DB 77 AATTCCAGCTGTTAATAATAGATGTGCAGAGGGCTTGAGAGTCTGCGGCTTTCTTGCTTG 136

QY 6484 AGGAG 6488  
DB 137 AGGAG 141

## RESULT 4

US-09-791-211-10  
Sequence 10, Application US/09791211  
Patent No. 6448080  
GENERAL INFORMATION:  
APPLICANT: Donna T. Ward  
TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION  
FILE REFERENCE: RTS-0205  
CURRENT APPLICATION NUMBER: US/09/791,211  
CURRENT FILING DATE: 2001-02-23  
NUMBER OF SEQ ID NOS: 90  
SEQ ID NO 10  
LENGTH: 98844  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 24962  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 64383  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 65468  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 65469  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 65470  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 65471  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 87130  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 89049  
OTHER INFORMATION: unknown  
OTHER INFORMATION: unknown  
US-09-791-211-10

Query Match 1.5%; Score 123.4; DB 4; Length 98844;  
Best Local Similarity 70.0%; Pred. No. 2.7e-22;  
Matches 166; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 3107 CCCCTCATCTTGTGAGGCTCCATTTGCCATTCACACACTTACATCTATGTGTACAC 3166  
DB 16525 CCACCATTCATCCGAGCCTCTGTATATCACCATTCTTACTTACTCTCATGACATCAA 16584

DB 16585 ATTTTGTAGCTCCGACATATAGTAGAGACATGGAGATATTGCTTCCCTCGCTGGTTT 16644

QY 3227 GTTACTTATGATATAGGCCCCAGTTTATCTAGGCTGTGCAAAAGGATGATTCA 3286  
DB 16645 ATTTACTTAAAGAGTCTTCCAGTTCATTCATGTTCTGCAAAATGACAGGATTCA 16704

QY 3287 TTCTTTTATGAGGTATGTTCTTCCCAATTAGATTAAGACACTCGACTTGTCTC 3343  
DB 16705 TTCTTTTATGAGGTATTAATATTCCATTGTATATATACAACTTTCTTATC 16761

## RESULT 5

US-10-027-983-11/c  
Sequence 11, Application US/10027983  
Patent No. 6617162  
GENERAL INFORMATION:  
APPLICANT: Kenneth W. Dobie  
APPLICANT: Mark P. Roach  
TITLE OF INVENTION: ANTISENSE MODULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION  
FILE REFERENCE: RTS-0340  
CURRENT APPLICATION NUMBER: US/10/027,983  
CURRENT FILING DATE: 2001-12-18  
NUMBER OF SEQ ID NOS: 98  
SEQ ID NO 11  
LENGTH: 392000  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 137740  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 137742  
OTHER INFORMATION: unknown  
NAME/KEY: misc feature  
LOCATION: (138122)...(138221)  
OTHER INFORMATION: n = A,T,C or G  
NAME/KEY: unsure  
LOCATION: 14507  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 151967  
OTHER INFORMATION: unknown  
NAME/KEY: misc feature  
LOCATION: (151567)...(1542066)  
OTHER INFORMATION: n = A,T,C or G  
NAME/KEY: unsure  
LOCATION: 154217  
OTHER INFORMATION: unknown  
NAME/KEY: misc feature  
LOCATION: (164037)...(164136)  
OTHER INFORMATION: n = A,T,C or G  
NAME/KEY: misc feature  
LOCATION: (174657)...(174756)  
OTHER INFORMATION: n = A,T,C or G  
NAME/KEY: misc feature  
LOCATION: (186324)...(186323)  
OTHER INFORMATION: n = A,T,C or G  
NAME/KEY: misc feature  
LOCATION: (195242)...(195341)  
OTHER INFORMATION: n = A,T,C or G  
NAME/KEY: unsure  
LOCATION: 202703  
OTHER INFORMATION: unknown  
NAME/KEY: misc feature  
LOCATION: (202771)...(202870)  
OTHER INFORMATION: n = A,T,C or G  
NAME/KEY: misc feature  
LOCATION: (206246)...(215602)  
OTHER INFORMATION: n = A,T,C or G  
NAME/KEY: misc feature  
LOCATION: (218126)...(218225)

```

OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc_feature
LOCATION: (220360)...(220459)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc_feature
LOCATION: (222717)...(222816)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc_feature
LOCATION: (223561)...(224080)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc_feature
LOCATION: (227487)...(227586)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc_feature
LOCATION: (230157)...(230256)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc_feature
LOCATION: (232299)...(232398)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc_feature
LOCATION: (236552)...(236651)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc_feature
LOCATION: (238789)...(248788)
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: exon
LOCATION: (118288)...(119101)
OTHER INFORMATION: exon 1C
NAME/KEY: exon:intron junction
LOCATION: (151129)...(151130)
OTHER INFORMATION: exon 5:intron 5
NAME/KEY: exon:intron junction
LOCATION: (299248)...(299249)
OTHER INFORMATION: exon 9:intron 9
NAME/KEY: exon:intron junction
LOCATION: (348578)...(348579)
OTHER INFORMATION: exon 10:intron 10
NAME/KEY: intron
LOCATION: (348579)...(381838)
OTHER INFORMATION: intron 10
NAME/KEY: intron:exon junction
LOCATION: (386185)...(386186)
OTHER INFORMATION: intron 11:exon 12
IS-10-027-983-11

```

	Query Match	1.5% ; Score 121.2 ; DB 4 ; Length 392000 ;
	Best Local Similarity	66.8% ; Pred. No. 2.9e-21 ;
	Matches 203 ; Conservative	0 ; Mismatches 98 ; Indels 3 ; Gaps 2 ;
QY	3042 TTGATCACCTGAATGATGATGCATTGTGACCCCTTAAGAAATTTCTACCAATCCGCTGACTT	3101
Db	99765 TTATACCACCGAGATTTAAGCCTAGTACCATT-AAGTATATTTTCTGATTCCTGCCCTC	99707
QY	3102 CTTGCCCCCTCATCTTCTG--AGGTCGATTGTGCATCTTCACACTCTACATTTATG	3159
Db	99706 CTTTCACCCCCTCACCCCTGTATAGGCCCGAGTGCTGTATTTCTCTATGATGTCACAG	99647
QY	3160 TGTACACATATTATTAGTGGTCTACTTATAAGTAAATACATGCAATATTGGCTTTCTGCTGT	3219
Db	99646 AATTCTAATCATGTAGCTCCCACTTAAGAATGAGAACATGCAATATTAGTTTCTGTTCC	99587
QY	3220 CTGCTCTTGTTTAACTTATGTAAATGGCCCCCAATTCATCTTAGCTGTGCAAAAAGCAT	3279
Db	99586 CTGATATTAGTTTGGTAAAGATAAAGGCTCCAGTCTTTCATAGTTTCTGAAAAAGCAAT	99522
QY	3280 GAATTCATTCTTTTAAATGGCGATGTTCTTTCCCAATTAGATAAAGAACTCGCACTT	3339
Db	99526 GATCTCATTTCTTTTGGTGGGTGATGATATCTATCATGTATATGACACATTTTCTT	99467
QY	3340 GCCTC 3343	
Db	99466 TATTC 99463	

```

1 RESULT 6
2 US-09-601-777-1
3 : Sequence 1, Application US/09601777
4 : Patent No. 6451848
5 :
6 : GENERAL INFORMATION:
7 :
8 : APPLICANT: Nakajima, Morowo
9 : APPLICANT: Funakubo, Minako
10 : TITLE OF INVENTION: Human heparanase polypeptide and cDNA
11 : FILE REFERENCE: 30384a
12 : CURRENT APPLICATION NUMBER: US/09/601,777
13 : CURRENT FILING DATE: 2000-08-07
14 : NUMBER OF SEQ ID NOS: 2
15 : SOFTWARE: PatentIn Ver. 2.1
16 : SEQ ID NO 1
17 : LENGTH: 3726
18 :
19 : TYPE: DNA
20 : ORGANISM: Human
21 US-09-601-777-1

```

	Query Match	1.5%; Score 119.6; DB 4	Length 3726;
	Best Local Similarity	65.4%; Pred. No. 3, 2e-22;	
	Matches 208; Conservative	0; Mismatches 104; Indels	Gaps 2;
QY	3039	GAATTCATCACCAGATGATGTGCATTGTACCCCTTAAGTAAATTTCTCACATCCGTGA	3098B
Db	3233	GAACCTGCACGTAGATGATGACATGTGCCAATAGATAATTTTCACTCATACCTT	3292Z
QY	3099	CTTCTTGCCCCCTCATCCTTTTAGGGTCCATTGTCCATCATTTCCACACTCTACATCAT	3158B
Db	3293	CCTTCGCCCTTCCTTCCTTGATGTCCACATCCCCTTAATACACAGTGTAGTCTT	3351B
QY	3159	GTFACACATTATTAGCTCCCTACTTAAAGATATACATGCATATTGTCTTCTGTG	3218B
Db	3352	GTGTACACACATACATACCTTCCACTTTAAGTATAGAACAATGCAATTTGGTTTTCCATT	3411B
QY	3219	TCTGTCTTGTTTAACTATAGATAATGCCCCCAGTTCTATCTAGCGTGTGCAAAAAGCA	3278B
Db	3412	CCTGATTAACCTCCCTTAGATTAACAGCCCCCAGTTCCTCCAAAGTGTGCAAAAAATACA	3471B
QY	3279	TGATTTCAATCTTTTATATGCGATGTCTCTTCCCAATTAATATAAAGAACCTGGCACT	3338B
Db	3472	-----TTATCTCTCTTAATGCGTAGTAATATGTCACTGTACATATATACACAACTTTTCT	3526B
QY	3339	TGCTCTACTCTCATTTTG	3356B
Db	3527	TTATCCACTTACAGTTG	3544B

```

RESULT 7
US-09-435-739-42
: Sequence 42: Application US/09435739
: Patent No. 5664105
: GENERAL INFORMATION:
: APPLICANT: Becker, Itis
: APPLICANT: Vlodavsky, Israel
: APPLICANT: Feinstein, Elena
: TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY
: TITLE OF INVENTION: EXPRESSION OF SAME IN GENETICALLY MODIFIED CELLS
: FILE REFERENCE: 00/20454
: CURRENT APPLICATION NUMBER: US/09/435,739
: CURRENT FILING DATE: 2001-06-05
: NUMBER OF SEQ ID NOS: 47
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 42
: LENGTH: 44848
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-435-739-42

```

Query Match	1.5%;	Score 119.6;	DB 4;	Length 44848;
Best Local Similarity	65.4%;	Pred. No. 1.8e-21;		
Matches 208;	Conservative 0;	Mismatches 104;	Indels 6;	Gaps 2;









Db 1407 GCCTGAGAAAA---AGTTAGCAGAGAAAAAGCATTAAC-GATTATTATTAATAAGTT 1353  
 QY 1290 TTACATGACATGAGAGCTCTCAGAAATGAAGACCCAAAGCCAGGGGAAACTGTCTGTT 1349  
 Db 1352 TTCTGTGACATGAGAGCTCTCAGAAATGAAGACCCAAAGCTTAAGAAAACTCTATTTT 1293  
 QY 1350 TTTTGTGAGGTTTCATGAGAGAGATGATGATGATGATGATGATGATGATGATGATG 1409  
 Db 1292 TAGGCTTAGATTCTATGAGAGACCATGAAATGC--GCAGAAATGTAGTGGGAAAAAG 1235  
 QY 1410 ATATGATCTAGTGTGAAGAGCTCAGGGGAAACACAGCAGGCGCTGTCTATTCAGATTC 1469  
 Db 1234 GAATGGTTAATAGTAATTAATGAGTGGGAAACACAGCAGGCGCTGTCTATTCAGATTC 1175  
 QY 1470 TTCTGATCTCTCTCTCTCTA 1490  
 Db 1174 TTCTGGCCCTATCTGTGGCCA 1154  
 RESULT 15  
 US-08-321-478-6/c  
 ; Sequence 6, Application US/08321478  
 ; Patent No. 5527677  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DEGUCHI, Takeo  
 ; APPLICANT: KINOSHITA, Moritoshi  
 ; APPLICANT: KATSURAGI, Kiyomori  
 ; APPLICANT: SHIN, Sadahito  
 ; TITLE OF INVENTION: HUMAN ARYLAMINE N-ACETYLTRANSFERASE  
 ; NUMBER OF SEQUENCES: 13  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Sughrue, Mion, Zimm, Macpeak & Seas  
 ; STREET: 2100 Pennsylvania Avenue, N.W.  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: United States  
 ; ZIP: 20037-3202  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/321,478  
 ; FILING DATE: 11-OCT-1994  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/038,667  
 ; FILING DATE: 23-MAR-1993  
 ; APPLICATION NUMBER: JP 64669/1992  
 ; FILING DATE: 23-MAR-1992  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202) 293-7060  
 ; TELEFAX: (202) 293-7860  
 ; TELEX: 6491103  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 6464 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 723..1595  
 ; FEATURE:  
 ; NAME/KEY: exon  
 ; LOCATION: 717..1936  
 ; FEATURE:  
 ; NAME/KEY: polyA signal  
 ; LOCATION: 1794..1799  
 ; FEATURE:

; NAME/KEY: polyA signal  
 ; LOCATION: 1800..1805  
 ; US-08-321-478-6  
 Query Match 1.4%; Score 115.8; DB 1; Length 6464;  
 Best Local Similarity 62.7%; Pred No. 5.3e-21;  
 Matches 180; Conservative 0; Mismatches 107; Indels 0; Gaps 0;  
 QY 3070 CCCTTAAGTAATTTCTCACCATCCGCTGACTTCTTGCCCCCTCATTCTTCTAGGCTCCA 3129  
 Db 5569 CACATTAATTTATTTTTCGCTTCTCTCTCTCCTCCTACTCTTCAACCCAGTAGGCCCA 5510  
 QY 3130 TTGTCATCATTCACACTGTACATGATGTGACATTAATTTAGCGCTTATAAG 3189  
 Db 5509 GTGTGAGTGTCCCTCTATGTGTCAATGTGTTCATCATTTTACTCCACTTAAAG 5450  
 QY 3190 TGATTAACATCAATATTGTCCTTCTGTGTCTGTCTGTTTACTATGATAATGCCCC 3249  
 Db 5449 TAAGAACATGTGATTTGGTTTTCGTTTCGATTTGCTTAAGATATGAGCCCTC 5390  
 QY 3250 CAGTTCTATCTAGGCTGCGCAAGAGCATGATTTGATCTTTTATGAGCTATGCTT 3309  
 Db 5389 CAGCTTCATCCATGTTTGTGCAAGAGCATGATCTATTTCTTATAGCTGACATGTA 5330  
 QY 3310 TCCCATTTAGATAAAGACACTGCACTTGCTCTTACTTCTATTG 3356  
 Db 5329 TTCCATGTGTATGTGACACATTTCTTTATCCAGTGTATCATTG 5283  
 Search completed: September 11, 2004, 20:14:21  
 Job time : 357 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 11, 2004, 16:46:54 ; Search time 2304 Seconds

(without alignments)  
17560.102 Million cell1 updates/sec

Title: US-09-939-209A-3\_COPY\_4110\_12150

Perfect score: 8041

Sequence: 1 caaatggaagctcctgggacac.....ctagatctcagctgaccc 8041

Scoring table:

Gapop 10.0 , Gapext 1.0

Searched: 3304383 seqs, 2515761380 residues

Total number of hits satisfying chosen parameters: 6608766

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCR\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8041	100.0	20300	US-09-939-209A-3	Sequence 3, Appli
2	600.8	7.5	666	US-10-027-632-68745	Sequence 68745, A
3	600.8	7.5	666	US-10-027-632-68745	Sequence 68745, A
4	598.2	7.4	615	US-10-027-632-68208	Sequence 68208, A
5	598.2	7.4	615	US-10-027-632-68209	Sequence 68209, A
6	598.2	7.4	615	US-10-027-632-68210	Sequence 68210, A
7	598.2	7.4	615	US-10-027-632-68211	Sequence 68211, A
8	598.2	7.4	615	US-10-027-632-68742	Sequence 68742, A
9	598.2	7.4	615	US-10-027-632-68743	Sequence 68743, A
10	598.2	7.4	615	US-10-027-632-68744	Sequence 68744, A
11	598.2	7.4	615	US-10-027-632-70068	Sequence 70068, A
12	598.2	7.4	615	US-10-027-632-70069	Sequence 70069, A
13	598.2	7.4	615	US-10-027-632-70070	Sequence 70070, A
14	598.2	7.4	615	US-10-027-632-70071	Sequence 70071, A

15	598.2	7.4	615	US-10-027-632-295009	Sequence 295009, A
16	598.2	7.4	615	US-10-027-632-295010	Sequence 295010, A
17	598.2	7.4	615	US-10-027-632-311585	Sequence 311585, A
18	598.2	7.4	615	US-10-027-632-311586	Sequence 311586, A
19	598.2	7.4	615	US-10-027-632-68208	Sequence 68208, A
20	598.2	7.4	615	US-10-027-632-68209	Sequence 68209, A
21	598.2	7.4	615	US-10-027-632-68210	Sequence 68210, A
22	598.2	7.4	615	US-10-027-632-68211	Sequence 68211, A
23	598.2	7.4	615	US-10-027-632-68742	Sequence 68742, A
24	598.2	7.4	615	US-10-027-632-68743	Sequence 68743, A
25	598.2	7.4	615	US-10-027-632-68744	Sequence 68744, A
26	598.2	7.4	615	US-10-027-632-70068	Sequence 70068, A
27	598.2	7.4	615	US-10-027-632-70069	Sequence 70069, A
28	598.2	7.4	615	US-10-027-632-70070	Sequence 70070, A
29	598.2	7.4	615	US-10-027-632-295009	Sequence 295009, A
30	598.2	7.4	615	US-10-027-632-295010	Sequence 295010, A
31	598.2	7.4	615	US-10-027-632-311585	Sequence 311585, A
32	598.2	7.4	615	US-10-027-632-311586	Sequence 311586, A
33	598.2	7.4	615	US-10-027-632-131556	Sequence 131556, A
34	277.6	3.5	2752	US-10-198-846-13843	Sequence 13843, A
35	276	3.4	1531	US-10-198-846-9863	Sequence 9863, A
36	274.4	3.4	3165	US-10-044-090-113	Sequence 113, A
37	274.4	3.4	3167	US-09-918-624B-6	Sequence 6, Appli
38	253.6	3.2	1151	US-10-027-632-266349	Sequence 266349, A
39	253.6	3.2	1151	US-10-027-632-266349	Sequence 266349, A
40	253.2	3.1	1151	US-10-027-632-266350	Sequence 266350, A
41	253.2	3.1	1151	US-10-027-632-266350	Sequence 266350, A
42	233.8	2.9	576	US-10-027-632-288758	Sequence 288758, A
43	233.8	2.9	576	US-10-027-632-288758	Sequence 288758, A
44	233.4	2.9	576	US-10-027-632-288755	Sequence 288755, A
45	233.4	2.9	576	US-10-027-632-288756	Sequence 288756, A

## ALIGNMENTS

RESULT 1  
US-09-939-209A-3  
Sequence 3, Application US/09939209A  
Publication No. US20030113721A1  
GENERAL INFORMATION:  
APPLICANT: LEVITT, PAT R.  
APPLICANT: MIRNICK, KAROLY  
APPLICANT: KODAVALI, MEKHATA CHONDARI  
APPLICANT: NINGAONKAR, VISHWALIT B.  
TITLE OF INVENTION: METHODS AND SYSTEMS FOR FACILITATING THE DIAGNOSIS AND TREATMENT  
TITLE OF INVENTION: SCHIZOPHRENIA  
FILE REFERENCE: 00-539-US  
CURRENT APPLICATION NUMBER: US/09/939, 209A  
CURRENT FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: 60/228,021  
PRIOR FILING DATE: 2000-08-24  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 20300  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: A genomic sequence containing RGS4 nucleic acid sequence and seq  
OTHER INFORMATION: encodes upstream and downstream to the RGS4 nucleic acid sequence  
US-09-939-209A-3  
Query Match 100.0%; Score 8041; DB 10; Length 20300;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 8041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAAATGAGCTCGGGACGCTTCAAACTAGAGGTGATTAAGTTTACCTTAATTACC 60  
Db 4110 CAAATGAGCTCGGGACGCTTCAAACTAGAGGTGATTAAGTTTACCTTAATTACC 4169  
QY 61 AAGATTCTTAGAGATCTATTAATGAATCCAGATCTGCTCTGTAAAGTTCAAGCA 120

Db 4170 AAGATTTCCATGAGGAGTCTATAATGGAATCCAGATCTGCTCTGTAAAGTTCAAGCA 4229  
Qy 121 CTTTCCATGACACCACTAGTTTCTTCCACCTGACATGGAATGAATCTTATGAA 180  
Db 4230 CTTTCCATGACACCACTAGTTTCTTCCACCTGACATGGAATGAATCTTATGAA 4289  
Qy 181 CTCTGTTTCTACTCTGGGCTAAATGTTGACAGAAAAAGATTTAATCTTTGGGTAAGGC 240  
Db 4290 CTCTGTTTCTACTCTGGGCTAAATGTTGACAGAAAAAGATTTAATCTTTGGGTAAGGC 4349  
Qy 241 TATTTGGGTTTCTCTACTCTTGGGAAACAAGTTTCTTCCCTGGCTAATTAAGT 300  
Db 4350 TATTTGGGTTTCTCTACTCTTGGGAAACAAGTTTCTTCCCTGGCTAATTAAGT 4409  
Qy 301 GTGATATTGTTCTTCCAGGAAATCAGTATGATCATCCTGCTGCTATCAATGTCAGG 360  
Db 4410 GTGATATTGTTCTTCCAGGAAATCAGTATGATCATCCTGCTGCTATCAATGTCAGG 4469  
Qy 361 TTGAGTTTCTGATTTTATTCATGTCCTCAAAAGCTTGTGCAAAAGATTGACACATT 420  
Db 4470 TTGAGTTTCTGATTTTATTCATGTCCTCAAAAGCTTGTGCAAAAGATTGACACATT 4529  
Qy 421 TCCCAAAAGTAACATACATCTGGGAAAGTCCCTGTTTACCTGCTGCTATCAAGATCTCC 480  
Db 4530 TCCCAAAAGTAACATACATCTGGGAAAGTCCCTGTTTACCTGCTGCTATCAAGATCTCC 4889  
Qy 481 AGCCCAATCTTCTGTTTATAGTCTTAAATCAATACTGAATCTCATTTAGTCTA 540  
Db 4590 AGCCCAATCTTCTGTTTATAGTCTTAAATCAATACTGAATCTCATTTAGTCTA 4649  
Qy 541 GGCCATTTAGTAAATTAAGAAAGAGGAGGCTTGTGACATCTGAGAGAAATGTC 600  
Db 4650 GGCCATTTAGTAAATTAAGAAAGAGGAGGCTTGTGACATCTGAGAGAAATGTC 4709  
Qy 601 ATCTGAAGTGTGCAAGCAGCCTGGGCTGAGGCTTGGCTATCTGCTGCTGCTGCTG 660  
Db 4710 ATCTGAAGTGTGCAAGCAGCCTGGGCTGAGGCTTGGCTATCTGCTGCTGCTGCTG 4769  
Qy 661 AGATAGAGTGGCCCATTTAACAATCTGTGTGAATCTTAAGAACGCAAGCTGCTGCAATGA 720  
Db 4770 AGATAGAGTGGCCCATTTAACAATCTGTGTGAATCTTAAGAACGCAAGCTGCTGCAATGA 4829  
Qy 721 CTTATTTCCCTCATTTGTAACCTCAATCTTGAATGTTTCTTTGTTGTTGCTGCTG 780  
Db 4830 CTTATTTCCCTCATTTGTAACCTCAATCTTGAATGTTTCTTTGTTGTTGCTGCTG 4889  
Qy 781 AGCAAAAGTTGAAAAATTTGAATTTAATCTAGAGACACATAGTTTCAATCTGCTG 840  
Db 4890 AGCAAAAGTTGAAAAATTTGAATTTAATCTAGAGACACATAGTTTCAATCTGCTG 4949  
Qy 841 TGTGCTGAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
Db 4950 TGTGCTGAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5009  
Qy 901 ACCTTCCAGATGTAGCATTTTACAATTTAAGAACTTAAATAGCCTTCAAACTTTTG 960  
Db 5010 ACCTTCCAGATGTAGCATTTTACAATTTAAGAACTTAAATAGCCTTCAAACTTTTG 5069  
Qy 961 CCAGTTTCTCTGATATCCAAATCTATTTTACTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
Db 5070 CCAGTTTCTCTGATATCCAAATCTATTTTACTGCTGCTGCTGCTGCTGCTGCTGCTG 5129  
Qy 1021 TGTCTAACCTGATGGCTTAAGTACTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
Db 5130 TGTCTAACCTGATGGCTTAAGTACTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5189  
Qy 1081 AATTCGAGTCTCAAGATTAACAAGAAACCTTTAATGTTCTGAGCATATCTGCTGCTG 1140  
Db 5190 AATTCGAGTCTCAAGATTAACAAGAAACCTTTAATGTTCTGAGCATATCTGCTGCTG 5249  
Qy 1141 CATTTATCTCTCTCTCTCTCAAGCTGACAGTGGGCTGAAACGTTGATATCATTTTC 1200

Db 5250 CATTTATCTCTCTCTCTCTCTCAAGCTGACAGTGGGCTGAAACGTTGATATCATTTTC 5309  
Qy 1201 TCTTCATCATTAAGGCTGGCAACCAAACTCTCTTATTTAAAGCAGCTTAATTAAGCA 1260  
Db 5310 TCTTCATCATTAAGGCTGGCAACCAAACTCTCTTATTTAAAGCAGCTTAATTAAGCA 5369  
Qy 1261 AAACCTTAACAATTTATTTAATCAAAAGTTTACATGACATGAGGCTTTCAGAAATGAAG 1320  
Db 5370 AAACCTTAACAATTTATTTAATCAAAAGTTTACATGACATGAGGCTTTCAGAAATGAAG 5429  
Qy 1321 ACCCAAAAGCCGAGGGAACCTGCTGTTTTTTTGTGAGGTTTCATGAAGATGATA 1380  
Db 5430 ACCCAAAAGCCGAGGGAACCTGCTGTTTTTTTGTGAGGTTTCATGAAGATGATA 5489  
Qy 1381 GCATGAGCCATGTGATTTAGACAAAGATATGATCTAGGTAAAGCACTCAGGGGGA 1440  
Db 5490 GCATGAGCCATGTGATTTAGACAAAGATATGATCTAGGTAAAGCACTCAGGGGGA 5549  
Qy 1441 AACACAGAGGCTGTCTATTTACAGATTTCTTTGATCTCTCTCTCTATGATAGCAT 1500  
Db 5550 AACACAGAGGCTGTCTATTTACAGATTTCTTTGATCTCTCTCTCTATGATAGCAT 5609  
Qy 1501 TCTTTCCTCTGAGTAAGGGGACAGACCTTCTTCAATGAGGGCTTCAAGGAGAGAGG 1560  
Db 5610 TCTTTCCTCTGAGTAAGGGGACAGACCTTCTTCAATGAGGGCTTCAAGGAGAGAGG 5669  
Qy 1561 AGAAAGTGCCTTTTATGATTTTAAGCTTGTGCTTGGGAAAGAGAGTTCTATTTCTAT 1620  
Db 5670 AGAAAGTGCCTTTTATGATTTTAAGCTTGTGCTTGGGAAAGAGAGTTCTATTTCTAT 5729  
Qy 1621 GACCCATTTGGGGAAGAAGAAATCTGCTTCTGATCTGCTTCAATGAAGAAAGAAGA 1680  
Db 5730 GACCCATTTGGGGAAGAAGAAATCTGCTTCTGATCTGCTTCAATGAAGAAAGAAGA 5789  
Qy 1681 GTAAGAGCAGAGGAGGAGATGTGACAAAGAGACTTGGCTGCTGAGAGGCTTCC 1740  
Db 5790 GTAAGAGCAGAGGAGGAGATGTGACAAAGAGACTTGGCTGCTGAGAGGCTTCC 5849  
Qy 1741 GCTCTCCTTTTATGTTCAAGTACTTCTTGAATACCAAGACATATCTTTGGCATTTGCT 1800  
Db 5850 GCTCTCCTTTTATGTTCAAGTACTTCTTGAATACCAAGACATATCTTTGGCATTTGCT 5909  
Qy 1801 TTTCTGAGCTCTAACATGCAATCATGCTAAATCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860  
Db 5910 TTTCTGAGCTCTAACATGCAATCATGCTAAATCTGCTGCTGCTGCTGCTGCTGCTGCTG 5969  
Qy 1861 TGTCTTTATTTCTTTATGTTGTGATGAGCATGAGGTCAATAGCAAAAGCCTGAGTCC 1920  
Db 5970 TGTCTTTATTTCTTTATGTTGTGATGAGCATGAGGTCAATAGCAAAAGCCTGAGTCC 6029  
Qy 1921 ACTGCTGAGCTGGCATTAATTAATCACTTAATCAATGCTGAGTAAAGTAC 1980  
Db 6030 ACTGCTGAGCTGGCATTAATTAATCACTTAATCAATGCTGAGTAAAGTAC 6089  
Qy 1981 TTAAGTCTCTATGCTTCAATCTGTAATTAAGAAATCAATGAAGAACATTTCTGAGATG 2040  
Db 6090 TTAAGTCTCTATGCTTCAATCTGTAATTAAGAAATCAATGAAGAACATTTCTGAGATG 6149  
Qy 2041 GATCATGAGAAATTAAGTAAATTAATGAGCAATAGGCTTAAACGAGTCCCTGCTGCTG 2100  
Db 6150 GATCATGAGAAATTAAGTAAATTAATGAGCAATAGGCTTAAACGAGTCCCTGCTGCTG 6209  
Qy 2101 TAGTGACAGATTAATATCATCTATTTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160  
Db 6210 TAGTGACAGATTAATATCATCTATTTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6269  
Qy 2161 TTGTACAAAGACTATCATCTATTTATCAATTTAAGGCTCTCTTGTGAAAAACAATCCA 2220  
Db 6270 TTGTACAAAGACTATCATCTATTTATCAATTTAAGGCTCTCTTGTGAAAAACAATCCA 6329  
Qy 2221 GATTCAAAAAGATTGAGTTTAACTCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280  
Db 6330 GATTCAAAAAGATTGAGTTTAACTCTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6389

QY	2281	ATCTTCACTGACATCATGATAGTCTTGAAAAGTCAACTTCAAGTGAAGACATA	2340	QY	3361	ACTAATTCCTAGGCTCTTGCAATGCTTTCTCTCTCAACCCATCAATCTCATTTTGA	3420
DB	6390	ATCTTCACTGACATCATGATAGTCTTGAAAAGTCAACTTCAAGTGAAGACATA	6449	DB	7470	ACTAATTCCTAGGCTCTTGCAATGCTTTCTCTCTCAACCCATCAATCTCATTTTGA	7529
QY	2341	ACTAACCATTCTTTCTCATCAAGCTTAATAATGAATGCAATGATGATGCA	2400	QY	3421	TACACCTCTTCAAGAGGCTTTCTGACACCTTGGCTGAATTAAGCCCTCACATCT	3480
DB	6450	ACTAACCATTCTTTCTCATCAAGCTTAATAATGAATGCAATGATGATGCA	6509	DB	7530	TACACCTCTTCAAGAGGCTTTCTGACACCTTGGCTGAATTAAGCCCTCACATCT	7589
QY	2401	TTCAAGACCGCTGATCCTTGTTCACTAAGTCACTGTTCCAAATATCATATGATG	2460	QY	3481	GATTACTCTGACACATCACTGCCATTTTATTCATGTAAGGTCAAAATCTGGAA	3540
DB	6510	TTCAAGACCGCTGATCCTTGTTCACTAAGTCACTGTTCCAAATATCATATGATG	6569	DB	7590	GATTACTCTGACACATCACTGCCATTTTATTCATGTAAGGTCAAAATCTGGAA	7649
QY	2461	ACATTGAGAGCTTACTATATATATATATATATATATATATATATATATATATAT	2520	QY	3541	CACCTGATTTGTTATTTTGTGACTCTCTCTACTGATGAAAACTCTACTGACGGAG	3600
DB	6570	ACATTGAGAGCTTACTATATATATATATATATATATATATATATATATATATAT	6629	DB	7650	CACCTGATTTGTTATTTTGTGACTCTCTCTACTGATGAAAACTCTACTGACGGAG	7709
QY	2521	CTGCTAATCTGCTAATCTGGTCTCCCTGCTTCCACAATCTGCCCATCTCATCTTT	2580	QY	3601	ATTATATCTGCTGATCAGGTACTGCTTCAACACACCTGATACAGATGAGTGTCA	3660
DB	6630	CTGCTAATCTGCTAATCTGGTCTCCCTGCTTCCACAATCTGCCCATCTCATCTTT	6689	DB	7710	ATTATATCTGCTGATCAGGTACTGCTTCAACACACCTGATACAGATGAGTGTCA	7769
QY	2581	CTTTCACAGAGTCAAGATGATGATGAGACCCCTCTGCTCTGTTCTTCTTCATGG	2640	QY	3661	AAAGATATTTCTTAAACAATGAACAATTAATAATTAATTAATTAATTAATTAATTA	3720
DB	6690	CTTTCACAGAGTCAAGATGATGATGAGACCCCTCTGCTCTGTTCTTCTTCATGG	6749	DB	7770	AAAGATATTTCTTAAACAATGAACAATTAATAATTAATTAATTAATTAATTAATTA	7829
QY	2641	ATTTCACATGCACTGTATTAAGTCCAGCTCTTGACCAAGCTTACAAATCTTGACAG	2700	QY	3721	CCACATCAGAGATCATTCAGGATGACAAATCATAGATTAACAGAAATTTGATGCTTG	3780
DB	6750	ATTTCACATGCACTGTATTAAGTCCAGCTCTTGACCAAGCTTACAAATCTTGACAG	6809	DB	7830	CCACATCAGAGATCATTCAGGATGACAAATCATAGATTAACAGAAATTTGATGCTTG	7889
QY	2701	ATCTATGCTTACTTTCATCTCTTTATGCTACTTTCATCTGTTCTCAATCTCTCA	2760	QY	3781	TGCATATCAGAGAAAGAGGTGAGAGTGTCAAGGTATCATGATGTACAGTCTCTGCC	3840
DB	6810	ATCTATGCTTACTTTCATCTCTTTATGCTACTTTCATCTGTTCTCAATCTCTCA	6869	DB	7890	TGCATATCAGAGAAAGAGGTGAGAGTGTCAAGGTATCATGATGTACAGTCTCTGCC	7949
QY	2761	GCTATGCTGCCCCCTCTGTTCTTCCCATTTTATATTTTAAATTTGATATATA	2820	QY	3841	TTCAACAACCTTTTAAAGTCTTAAACCTCTTAAGAAAGCAAGATTTAATAGTCAAG	3900
DB	6870	GCTATGCTGCCCCCTCTGTTCTTCCCATTTTATATTTTAAATTTGATATATA	6929	DB	7950	TTCAACAACCTTTTAAAGTCTTAAACCTCTTAAGAAAGCAAGATTTAATAGTCAAG	8009
QY	2821	TTTATGAGTATTAAGTGAATCTTTTATAGATGCAATGATGATGATGATGATGATG	2880	QY	3901	TTCAACAACCTTTTAAAGTCTTAAACCTCTTAAGAAAGCAAGATTTAATAGTCAAG	3960
DB	6930	TTTATGAGTATTAAGTGAATCTTTTATAGATGCAATGATGATGATGATGATGATG	6989	DB	8010	TTCAACAACCTTTTAAAGTCTTAAACCTCTTAAGAAAGCAAGATTTAATAGTCAAG	8069
QY	2881	CTTTAGGATATTCATCACTGAAATGATGATCATTTGATCCCTTAAGTAAATTTCTGACA	2940	QY	3961	AAAGTGAACAATGAAATGCTTGAACAGATGATGATGATGATGATGATGATGATG	4020
DB	6990	CTTTAGGATATTCATCACTGAAATGATGATCATTTGATCCCTTAAGTAAATTTCTGACA	7049	DB	8070	AAAGTGAACAATGAAATGCTTGAACAGATGATGATGATGATGATGATGATGATG	8129
QY	2941	TCCGCTGATCTTGGCCCCCTGGGATTTCACTCACTGAATGATGATGATGATGATG	3000	QY	4021	GTTAGTAAATATGCTTGAAGACAGAGCTCTGCTTCAAGAGCAAGAAATTTGGG	4080
DB	7050	TCCGCTGATCTTGGCCCCCTGGGATTTCACTCACTGAATGATGATGATGATGATG	7109	DB	8130	GTTAGTAAATATGCTTGAAGACAGAGCTCTGCTTCAAGAGCAAGAAATTTGGG	8189
QY	3001	AGTAAATTTCTACCATCCGCTGATCTTGCCCCCTGGGATTTCACTCACTGAATGATG	3060	QY	4081	CAAGTGAACAATGAAATGCTTGAACAGATGATGATGATGATGATGATGATGATG	4140
DB	7110	AGTAAATTTCTACCATCCGCTGATCTTGCCCCCTGGGATTTCACTCACTGAATGATG	7169	DB	8190	CAAGTGAACAATGAAATGCTTGAACAGATGATGATGATGATGATGATGATGATG	8249
QY	3061	GCATTGTAACCCCTTAAGTAAATTTCTACCATCCGCTGATCTTGCCCCCTCAATCTTCT	3120	QY	4141	GTATTAATGAATGAATTAATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA	4200
DB	7170	GCATTGTAACCCCTTAAGTAAATTTCTACCATCCGCTGATCTTGCCCCCTCAATCTTCT	7229	DB	8250	GTATTAATGAATGAATTAATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA	8309
QY	3121	GAGGCTGCAATTTGCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT	3180	QY	4201	TACACAGTTTACCTTATTTATTAAGACTGTTGAGTGAAGTGAAGTGAAGTGAAGTGA	4260
DB	7230	GAGGCTGCAATTTGCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT	7289	DB	8310	TACACAGTTTACCTTATTTATTAAGACTGTTGAGTGAAGTGAAGTGAAGTGAAGTGA	8369
QY	3181	ACTTAATGATGAATCAATGATATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3240	QY	4261	ACATTAATGATGAATCAATGATATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4320
DB	7290	ACTTAATGATGAATCAATGATATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	7349	DB	8370	ACATTAATGATGAATCAATGATATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	8429
QY	3241	AATGGCCCCCATTTTATCTAGGCTGCTGCAAAAGGATGATTTTATTTATGAGC	3300	QY	4321	CCTGAGCAAGTCACTTCTCTTCTGAGTCTTATTTCTTTTGGTGAATGAGTGT	4380
DB	7350	AATGGCCCCCATTTTATCTAGGCTGCTGCAAAAGGATGATTTTATTTATGAGC	7409	DB	8430	CCTGAGCAAGTCACTTCTCTTCTGAGTCTTATTTCTTTTGGTGAATGAGTGT	8489
QY	3301	TATGTTCTTTCCCAATTTAGATTAAGAACTGCACTGCTTACTTCTATTTGGAAT	3360	QY	4381	TAGGCTGATTTGCTTTGAAGTCCATTTTGTCTTAAAGTCCCATCTATGAGTGAAT	4440
DB	7410	TATGTTCTTTCCCAATTTAGATTAAGAACTGCACTGCTTACTTCTATTTGGAAT	7469	DB	8490	TAGGCTGATTTGCTTTGAAGTCCATTTTGTCTTAAAGTCCCATCTATGAGTGAAT	8549
QY				QY	4441	TATATTAATCATGACAAATCAGGTTCTCTATTTCAAGTCAAGACATAAACTTTT	4500

Db 8550 TAAATTTAACTGACGCAATCAGGCTTCTTATTTCTAAGGCAAGCATMAAACTTTT 8609  
Qy 4501 ATTGGAAATTTGAGCATCAGTAAATCTTTTGGGTACTGCTATGTTCTGTAATCA 4560  
Db 8610 ATTGGAAATTTGAGCATCAGTAAATCTTTTGGGTACTGCTATGTTCTGTAATCA 8669  
Qy 4561 ATCTATTGAGTATCTCTTTTAGTGGCCAGGTAAACAAAGAGCCATGTCCTTC 4620  
Db 8670 ATCTATTGAGTATCTCTTTTAGTGGCCAGGTAAACAAAGAGCCATGTCCTTC 8729  
Qy 4621 TTGAGTGAACCTTCTTCCCTTTATTTAGTGTGACCTTTAATGAGTCTGACCTGA 4680  
Db 8730 TTGAGTGAACCTTCTTCCCTTTAATTTAGTGTGACCTTTAATGAGTCTGACCTGA 8789  
Qy 4681 TTCAATTCCTGCTCCATCTTCTTGGTGTGAGGACCTTCTAGTTTCAATGACCTTC 4740  
Db 8790 TTCAATTCCTGCTCCATCTTCTTGGTGTGAGGACCTTCTAGTTTCAATGACCTTC 8849  
Qy 4741 AGTTCTTCCACACCCATCAAGATGCTGTCAACTTATTGTTCTATGTTAAT 4800  
Db 8850 AGTTCTTCCACACCCATCAAGATGCTGTCAACTTATTGTTCTATGTTAAT 8809  
Qy 4801 TCAAGAAAAGTGGCCAGTACTATCCATMAATGCCCTTTATGGGCGGTAGAC 4860  
Db 8910 TCAAGAAAAGTGGCCAGTACTATCCATMAATGCCCTTTATGGGCGGTAGAC 8969  
Qy 4861 TTTTCTCTAATTTAATGCACTTCTGTGATATATGTCCTTCACACTTTAATTT 4920  
Db 8970 TTTTCTCTAATTTAATGCACTTCTGTGATATATGTCCTTCACACTTTAATTT 9029  
Qy 4921 GTCTGCTCTCTCTGCTGCTAGTCAAGAACCTGTGGTAGGGGGAGTGTGGA 4980  
Db 9030 GTCTGCTCTCTCTGCTGCTAGTCAAGAACCTGTGGTAGGGGGAGTGTGGA 9089  
Qy 4981 TGTTCAAGGACAGTATGGGTAGGGCCACATATGGGCAATGCTTTGTGCCATTTCT 5040  
Db 9090 TGTTCAAGGACAGTATGGGTAGGGCCACATATGGGCAATGCTTTGTGCCATTTCT 9149  
Qy 5041 ATATTTTGGTATTTGCACTCTCAGTGGAACCACTATTTTCTGCTTCCACCTAA 5100  
Db 9150 ATATTTTGGTATTTGCACTCTCAGTGGAACCACTATTTTCTGCTTCCACCTAA 9209  
Qy 5101 CTATTTGATGCTCTGTTTCTTATATATMAAGTATAGCTCACTGATGCTATGACGA 5160  
Db 9210 CTATTTGATGCTCTGTTTCTTATATATMAAGTATAGCTCACTGATGCTATGACGA 9269  
Qy 5161 ACCATGCTCTTCTMAATGAAAGCTTTTGGTCAAGTCTAGCAATTAATTCCTTCTT 5220  
Db 9270 ACCATGCTCTTCTMAATGAAAGCTTTTGGTCAAGTCTAGCAATTAATTCCTTCTT 9329  
Qy 5221 CCACTTATGCTTCTCTGTACTGCTGTAGGTATTTGTTATGCTATAGATG 5280  
Db 9330 CCACTTATGCTTCTCTGTACTGCTGTAGGTATTTGTTATGCTATAGATG 9389  
Qy 5281 GAAACCTGATGATCTGTGATGACGAGCAATTCAGTTATATATTTGATGTAAG 5340  
Db 9390 GAAACCTGATGATCTGTGATGACGAGCAATTCAGTTATATATTTGATGTAAG 9449  
Qy 5341 ATACTGATGCTGTAAGTGTGAAACACCTATAGGGCTTAAATTTCTTCAATCACTT 5400  
Db 9450 ATACTGATGCTGTAAGTGTGAAACACCTATAGGGCTTAAATTTCTTCAATCACTT 9509  
Qy 5401 GAAAGCTGTGATACGAAAGCAATCAATATTTTCTGCTGCTCTCTCTTCTCT 5460  
Db 9510 GAAAGCTGTGATACGAAAGCAATCAATATTTTCTGCTGCTCTCTCTCTCTCT 9569  
Qy 5461 CTACATCTCTCTTCTTATCTTTTGAATATCACTGTTGAGACTTGAATTCATAGA 5520  
Db 9570 CTACATCTCTCTTCTTATCTTTTGAATATCACTGTTGAGACTTGAATTCATAGA 9629  
Qy 5521 CATAAACCAATTTGATATAGAAATGCTGTATATTTGCTCATCTACCTCTTGTG 5580

Db 9630 CATAAACCAATTTGATATAGAAATGCTGTATATTTGCTCATCTACCTCTTGTG 9689  
Qy 5581 TCCCTGAGCTCCGGTTAGACTTTTACAGAGCAGCATGTGAAGAACTGTCA 5640  
Db 9690 TCCCTGAGCTCCGGTTAGACTTTTACAGAGCAGCATGTGAAGAACTGTCA 9749  
Qy 5641 GTGCTAGCTGAATTTCTGTTTACCAAGATTTCTAGAAAAGTATCTCAGTGTG 5700  
Db 9750 GTGCTAGCTGAATTTCTGTTTACCAAGATTTCTAGAAAAGTATCTCAGTGTG 9809  
Qy 5701 ATTACAGATATAGCAATCTATTTTCTGAGGAGTGTGTATGCTGCGGGCTTATA 5760  
Db 9810 ATTACAGATATAGCAATCTATTTTCTGAGGAGTGTGTATGCTGCGGGCTTATA 9869  
Qy 5761 ACTGCTGTCACTCAGCTATTTCTCTCAGCTTCTGTTGCAATCAACCAAGCACT 5820  
Db 9870 ACTGCTGTCACTCAGCTATTTCTCTCAGCTTCTGTTGCAATCAACCAAGCACT 9929  
Qy 5821 TCCGCAATCACTGCGTGAAGAGATGATCTGCAAGTCCCTTTGAAATGCTGAGA 5880  
Db 9930 TCCGCAATCACTGCGTGAAGAGATGATCTGCAAGTCCCTTTGAAATGCTGAGA 9989  
Qy 5881 TCAGATCTTGGACCATGATATATATGATGCTTCAATCCAAAAGAGAAAGCATTTGGA 5940  
Db 9990 TCAGATCTTGGACCATGATATATATGATGCTTCAATCCAAAAGAGAAAGCATTTGGA 10049  
Qy 5941 GTCAGCTCTTAAGTACCTCAGAAATTCCTGCTGTACTTTCTCTCCAGAGACACTT 6000  
Db 10050 GTCAGCTCTTAAGTACCTCAGAAATTCCTGCTGTACTTTCTCTCCAGAGACACTT 10109  
Qy 6001 CTTGATATTTTCTTTTACAGCATATGATATATAATTTTGCAAGATTTGACAC 6060  
Db 10110 CTTGATATATTTTCTTTTACAGCATATGATATATAATTTTGCAAGATTTGACAC 10169  
Qy 6061 TTTTCTCTTCTTTCTAAGAAATTTCAAACTCTGTACATGCTGTGACATTTGATACATTT 6120  
Db 10170 TTTTCTCTTCTTTCTAAGAAATTTCAAACTCTGTACATGCTGTGTGACATTTGATACATTT 10229  
Qy 6121 TTTCCATATCCCTACTTTTCAAGAGATTTCTGCTGTGCTCACTTAACATTTGCTAT 6180  
Db 10220 TTTCCATATCCCTACTTTTCAAGAGATTTTCTGCTGTGCTCACTTAACATTTGCTAT 10289  
Qy 6181 GCGTCACTCTTTCTTCTCATCTCTTTCAGGGGCTGTGAGAGCAGAGGAGACAGAGA 6240  
Db 10290 GCGTCACTCTTTCTTCTCATCTCTTTCAGGGGCTGTGAGAGCAGAGGAGACAGAGA 10349  
Qy 6241 GCTGTACTGAGAGCGGTGCTGTGATTTGCTGTGAGAGGCTGTAGCTGGGCTTAAAGAG 6300  
Db 10350 GCTGTACTGAGAGCGGTGCTGTGATTTGCTGTGAGAGGCTGTAGCTGGGCTTAAAGAG 10409  
Qy 6301 ACCCTTACAGCTTTAGAGAGACGCTCAGAGATTTCTGACATATCTTTACCGAGAGA 6360  
Db 10410 ACCCTTACAGCTTTAGAGAGACGCTCAGAGATTTCTGACATATCTTTACCGAGAGA 10469  
Qy 6361 GAGGCAAGTACGCTCAAAACCGACACAGCTCTCTGCGGCAATTTCTTCTGCTT 6420  
Db 10470 GAGGCAAGTACGCTCAAAACCGACACAGCTCTCTGCGGCAATTTCTTCTGCTT 10529  
Qy 6421 GCGAATTCAGGCTTTAATATAGATGTGAAAGGCTTGCAGGTCTGCGGCTTCTTC 6480  
Db 10530 GCGAATTCAGGCTTTAATATAGATGTGAAAGGCTTGCAGGTCTGCGGCTTCTTC 10589  
Qy 6481 TTGAGAGGTATGATTTGCTTTACGCAATTAACATATTTAACTTTTGGCTAATCTTCTC 6540  
Db 10590 TTGAGAGGTATGATTTGCTTTACGCAATTAACATATTTAACTTTTGGCTAATCTTCTC 10649  
Qy 6541 AGTATTTAATGATGTTACTTACTTAACCTAGTCTGTGCAATTTAGAAACAGTGTGTGAG 6600  
Db 10650 AGTATTTAATGATGTTACTTACTTAACCTAGTCTGTGCAATTTAGAAACAGTGTGTGAG 10709  
Qy 6601 GAGAGCAGACTTTCTAATCTTCTCAGAGCTAGTATATGTAATTAAGCATGT 6660  
Db 10710 GAGAGCAGACTTTCTAATCTTCTCAGAGCTAGTATATGTAATTAAGCATGT 10769

```
QY 6661 GCTCCCAATTTCAGCCCTTATGTGTGTTGTGTGACCTCAGTTTGGAGACTGTTC 6720
DB 10770 GCTCCCAAAATTTCAGCCCTTATGTGTGTTGTGTGACCTCAGTTTGGAGACTGTTC 10829
QY 6721 TATCTTTAAGCAGGTCTAAGAAAGCTAGTTTAAATTAAAGCCGAGATGAGTTTGAG 6780
DB 10830 TATCTTTAAGCAGGTCTAAGAAAGCTAGTTTAAATTAAAGCCGAGATGAGTTTGAG 10889
QY 6781 GCTATGACAGTATCTGTAAATATCTCCATCTGTGATTACTACTGTATTGAGATCCC 6840
DB 10890 GCTATGACAGTATCTGTAAATATCTCCATCTGTGATTACTACTACTGTATTGAGATCCC 10949
QY 6841 TGGAGTACATAGAAAGCCGCTCGGCTTCGATTCGATTAAGTACAGTGTGTTAGAGA 6900
DB 10950 TGGAGTACATAGAAAGCCGCTCGGCTTCGATTCGATTAAGTACAGTGTGTTAGAGA 11009
QY 6901 AAGTACCCCAAGATGAGGTTGAGCTCCATCATCAGAAAGCACTATGCTTTCGTTGAG 6960
DB 11010 AAGTACCCCAAGATGAGGTTGAGCTCCATCATCAGAAAGCACTATGCTTTCGTTGAG 11069
QY 6961 TGGTGCAGTAACTTCTCATCTCTATGTCTTATAGCAAAATGTTACATAGATATGAGT 7020
DB 11070 TGGTGCAGTAACTTCTCATCTCTATGTCTTATAGCAAAATGTTACATAGATATGAGT 11129
QY 7021 TTTAAAGCCAGATCTTCTTATCTCTCGCCCATCTCTATGTTCTTGAAGTGTCTCATAT 7080
DB 11130 TTTAAAGCCAGATCTTCTTATCTCTCGCCCATCTCTATGTTCTTGAAGTGTCTCATAT 11189
QY 7081 GAGTTTGGTGGAAATATGATCATTAACAATAGTAAATAGTTTGTGTAAGATCTCA 7140
DB 11190 GAGTTTGGTGGAAATATGATCATTAACAATAGTAAATAGTTTGTGTAAGATCTCA 11249
QY 7141 TCTTAAAGACATGTTGTTGTTAAATATCTCCCTGATTTTAAAGACCTTACAGACA 7200
DB 11250 TCTTAAAGACATGTTGTTGTTAAATATCTCCCTGATTTTAAAGACCTTACAGACA 11309
QY 7201 TACAGCTTATCATTTGTTTGTGTTGTTCAAAAAGGTAATAAGAAATCATTCAGAGA 7260
DB 11310 TACAGCTTATCATTTGTTTGTGTTGTTCAAAAAGGTAATAAGAAATCATTCAGAGA 11369
QY 7261 AAGATCAATATTTAGCCAGTGGAAATTAACAACAATAAGTGCATATGACTTCTA 7320
DB 11370 AAGATCAATATTTAGCCAGTGGAAATTAACAACAATAAGTGCATATGACTTCTA 11429
QY 7321 ATCTTGAGTCAGAAAGGTAAAGTCAACCTTAAGGTACTACCTGTTCTTATGCGAC 7380
DB 11430 ATCTTGAGTCAGAAAGGTAAAGTCAACCTTAAGGTACTACCTGTTCTTATGCGAC 11489
QY 7381 TGCAAATATGAATTTGCCCAATTTTATTTGGAAATTAATCTCAGAAACATATTTTTT 7440
DB 11490 TGCAAATATGAATTTGCCCAATTTTATTTGGAAATTAATCTCAGAAACATATTTTTT 11549
QY 7441 ATGTACTATTAACATTTTATCTTCCAAATATCTGTCAATTCAGAGATGAGATATCG 7500
DB 11550 ATGTACTATTAACATTTTATCTTCCAAATATCTGTCAATTCAGAGATGAGATATCG 11609
QY 7501 ATGCTTCTTTAAATGAGAGAGAGGCTCGAGAGAGATCTATGAATAAAGTTCCCT 7560
DB 11610 ATGCTTCTTTAAATGAGAGAGAGGCTCGAGAGAGATCTATGAATAAAGTTCCCT 11669
QY 7561 CTGACCTTCAGGCTTAATTTCTGAATGAGTGAAGGAATTAATCTCAACCTTCACTTA 7620
DB 11670 CTGACCTTCAGGCTTAATTTCTGAATGAGTGAAGGAATTAATCTCAACCTTCACTTA 11729
QY 7621 CTTCGATATGAATTCGACCTTACAAATAATACAGAGTGTGAGAGAAAGTTATGCTCTG 7680
DB 11730 CTTCGATATGAATTCGACCTTACAAATAATACAGAGTGTGAGAGAAAGTTATGCTCTG 11789
QY 7681 TAAATATTTTGGAAACAGATTAAGAAATATCTAGAGCTCTGCTCCCAAGAGTTAAGC 7740
DB 11790 TAAATATTTTGGAAACAGATTAAGAAATATCTAGAGCTCTGCTCCCAAGAGTTAAGC 11849
```

```
QY 7741 AGCTAATCTAAGAGAGTAAACTATATGTACAGAGATGAATCTCTTCTCTCTCTCT 7800
DB 11850 AGCTAATCTAAGAGAGTAAACTATATGTACAGAGATGAATCTCTTCTCTCTCTCT 11909
QY 7801 CAATTAATTTGGAATATCATCTAGTCCAAACATCTTTACACAGAGTCCGAGAG 7860
DB 11910 CAATTAATTTGGAATATCATCTAGTCCAAACATCTTTACACAGAGTCCGAGAG 11969
QY 7861 AGCCATTGCTCTTCAAGGTGCACATAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 7920
DB 11970 AGCCATTGCTCTTCAAGGTGCACATAGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 12029
QY 7921 CTGACTCCAGTAACCTTCTGAAGTATTTGTTTATTTTATTTATTTATTTATTAAGAA 7980
DB 12030 CTGACTCCAGTAACCTTCTGAAGTATTTGTTTATTTTATTTATTTATTTATTAAGAA 12089
QY 7981 TACTTGTAAAGACACTTACCCCTCGATGATTAATTAATCTTAGATTCAGGTGATC 8040
DB 12090 TACTTGTAAAGACACTTACCCCTCGATGATTAATTAATCTTAGATTCAGGTGATC 12149
QY 8041 C 8041
DB 12150 C 12150

RESULT 2
US-10-027-632-68745
; Sequence 68745, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827, 129
; CURRENT APPLICATION NUMBER: US/10/027, 632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218, 006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198, 676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193, 483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185, 218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167, 363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156, 358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146, 002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 68745
; LENGTH: 666
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-68745

Query Match 7.5%; Score 600.8; DB 13; Length 666;
Best Local Similarity 98.1%; Pred. No. 3.1e-134;
Matches 614; Conservative 4; Mismatches 6; Indels 2; Gaps 1;
```

Db	1	AGCGTTGGTGCAAAAGAAATTGGACACATTTCCCAAAAGATGAACACATACTGGGAAGTCCCTG	60
QY	453	TTTACCTTCCTGGTATPAGACATCCTCAGCCCATATCTTGGTTTGTAGCTCTAAAA	512
Db	61	TTTACCTTCCTGGTATPAGACATCCTCAGCCCATATCTTGGTTTGTAGCTCTAAAA	120
QY	513	TCAATTACTGAACCTCTCATTTGATGTCTAGGCCATTTAGTAAACATAAAGAGAGGGA	572
Db	121	TCAATTACTGAACCTCTCATTTGATGTCTAGGCCATTTAGTAAACATAAAGAGAGGGA	180
QY	573	GGCTTCGTGCAACTGAGAGAAATTTGATCTGAAGAGGTGCAAGCACAGCCTGGGGCTG	632
Db	181	GGCTTCGTGCAACTGAGAGAAATTTGATCTGAAGAGGTGCAAGCACAGCCTGGGGCTG	240
QY	633	AGCCTTGGCTTACATCTCTGCCAAGTGGAGGATCAGTGGCCCATTTAACCTTGGTGAAG	692
Db	241	AGCCTTGGCTTACATCTCTGCCAAGTGGAGGATCAGTGGCCCATTTAACCTTGGTGAAG	300
QY	693	CTPAAAGACGCAAGCCCTGCACACATGACTTAATTTCCCTGCATTTGATACCGTCAATCCT	752
Db	301	CTPAAAGACGCAAGCCCTGCACACATGACTTAATTTCCCTGCATTTGATACCGTCAATCCT	360
QY	753	TGAGAAAGTCTTTCTTTTGTCTCCCTGAGGAAGGTGGAAAAATTTGAAATTTACCTA	812
Db	361	TGAGAAAGTCTTTCTTTTGTCTCCCTGAGGAAGGTGGAAAAATTTGAAATTTACCTA	420
QY	813	GAGACCACACATAGTTCACATCTCTGCTGTGTGCTGTAATGTCTGCCCCCAGTAGGAAAC	872
Db	421	GAGACCACACATAGTTCACATCTCTGCTGTGTGCTGTAATGTCTGCCCCCAGTAGGAAAC	480
QY	873	AGTTCCTTCAAGCCTATTTGTCAACAATACCTTCCCAATGTATGACATTTTACAAATTAAAG	932
Db	481	AGTTCCTTCAAGCCTATTTGTCAACAATACCTTCCCAATGTATGACATTTTACAAATTAAAG	538

```

US-10-027-632-68208
; Sequence 68208, Application US/10027632
; Publication No. US20020198571A1
; GENERAL INFORMATION:
; APPLICANT: Mang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108627.125
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68208
; LENGTH: 615
; TYPE: DNA

```



ORGANISM: Human  
US-10-027-632-68208

Query Match 7.4%; Score 598.2; DB 13; Length 615;  
Best Local Similarity 98.7%; Pred. No. 1.2e-133;  
Matches 609; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

393 AAGCTTGGTGCAGAAATTGGACACATTTCCCAAAAGTAGACATACGTGGAGTCCCTG 452  
1 AAGCTTGGTGCAGAAATTGGACACATTTCCCAAAAGTAGACATACGTGGAGTCCCTG 60  
453 TTTACCTTCCCTGGATATACAGATCCTCCAGCCCATATCTTTGGTTTATGCTCTAAAA 512  
61 TTTACCTTCCCTGGATATACAGATCCTCCAGCCCATATCTTTGGTTTATGCTCTAAAA 120  
513 TCAATACTGAACCTCTCATGATGTCTAGGCACTTGTAGTAAACAATAAAGAGAGGGA 572  
121 TCAATACTGAACCTCTCATGATGTCTAGGCACTTGTAGTAAACAATAAAGAGAGGGA 180  
573 GGCTTCTGACAACTGAGAGGAAATTGTCATCTGAAGTGTGCAAGCAGAGCTGGGGCTG 632  
181 GGCTTCTGACAACTGAGAGGAAATTGTCATCTGAAGTGTGCAAGCAGAGCTGGGGCTG 240  
633 AGCCTTGGCTACATCTCCGCAAGTGGAGGATAGTGGCCCATTTAATCTGTGTAGAA 692  
241 AGCCTTGGCTACATCTCCGCAAGTGGAGGATAGTGGCCCATTTAATCTGTGTAGAA 300  
633 CTAAAGAAAGGAGAGGCTGCGCAATGACTTATTTCCCTGATTTAGTAAACCTGCAATCT 752  
301 CTAAAGAAAGGAGAGGCTGCGCAATGACTTATTTCCCTGATTTAGTAAACCTGCAATCT 360  
753 TGAGAAATGTTTCTTTTGTCTCCCTGAGCAAGTGTGAAAAATTGAAATTTACCTA 812  
361 TGAGAAATGTTTCTTTTGTCTCCCTGAGCAAGTGTGAAAAATTGAAATTTACCTA 420  
813 GAGACCAACATAGTTCATCTCTGCTGTGTGCTGATGCTGCCCCAGTAGAGAAAC 872  
421 GAGACCAACATAGTTCATCTCTGCTGTGTGCTGATGCTGCCCCAGTAGAGAAAC 480  
873 AGTCTCTTAAAGCCTATTTGCAACAATACCTTCCAGATGTTAGCATTTTAACTTAAG 932  
481 AGTCTCTTAAAGCCTATTTGCAACAATACCTTCCAGATGTTAGCATTTTAACTTAAG 538  
933 GAACCTTAAATAGCCTTCAAACTTTTGGCCAGTTCTGTATTCATTCATTTCTTTA 992  
539 GAACCTTAAATAGCCTTCAAACTTTTGGCCAGTTCTGTATTCATTCATTTCTTTA 598  
993 CTCTGCTCTCCAGCTT 1009  
599 CTCTGCTCTCCAGCTT 615

RESULT 5  
US-10-027-632-68209  
; Sequence 68209, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 68209  
LENGTH: 615  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-68209

Query Match 7.4%; Score 598.2; DB 13; Length 615;  
Best Local Similarity 98.7%; Pred. No. 1.2e-133;  
Matches 609; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

393 AAGCTTGGTGCAGAAATTGGACACATTTCCCAAAAGTAGACATACGTGGAGTCCCTG 452  
1 AAGCTTGGTGCAGAAATTGGACACATTTCCCAAAAGTAGACATACGTGGAGTCCCTG 60  
453 TTTACCTTCCCTGGATATACAGATCCTCCAGCCCATATCTTTGGTTTATGCTCTAAAA 512  
61 TTTACCTTCCCTGGATATACAGATCCTCCAGCCCATATCTTTGGTTTATGCTCTAAAA 120  
513 TCAATACTGAACCTCTCATGATGTCTAGGCACTTGTAGTAAACAATAAAGAGAGGGA 572  
121 TCAATACTGAACCTCTCATGATGTCTAGGCACTTGTAGTAAACAATAAAGAGAGGGA 180  
573 GGCTTCTGACAACTGAGAGGAAATTGTCATCTGAAGTGTGCAAGCAGAGCTGGGGCTG 632  
181 GGCTTCTGACAACTGAGAGGAAATTGTCATCTGAAGTGTGCAAGCAGAGCTGGGGCTG 240  
633 AGCCTTGGCTACATCTCCGCAAGTGGAGGATAGTGGCCCATTTAATCTGTGTAGAA 692  
241 AGCCTTGGCTACATCTCCGCAAGTGGAGGATAGTGGCCCATTTAATCTGTGTAGAA 300  
693 CTAAAGAAAGGAGAGGCTGCGCAATGACTTATTTCCCTGATTTAGTAAACCTGCAATCT 752  
301 CTAAAGAAAGGAGAGGCTGCGCAATGACTTATTTCCCTGATTTAGTAAACCTGCAATCT 360  
753 TGAGAAATGTTTCTTTTGTCTCCCTGAGCAAGTGTGAAAAATTGAAATTTACCTA 812  
361 TGAGAAATGTTTCTTTTGTCTCCCTGAGCAAGTGTGAAAAATTGAAATTTACCTA 420  
813 GAGACCAACATAGTTCATCTCTGCTGTGTGCTGATGCTGCCCCAGTAGAGAAAC 872  
421 GAGACCAACATAGTTCATCTCTGCTGTGTGCTGATGCTGCCCCAGTAGAGAAAC 480  
873 AGTCTCTTAAAGCCTATTTGCAACAATACCTTCCAGATGTTAGCATTTTAACTTAAG 932  
481 AGTCTCTTAAAGCCTATTTGCAACAATACCTTCCAGATGTTAGCATTTTAACTTAAG 538  
933 GAACCTTAAATAGCCTTCAAACTTTTGGCCAGTTCTGTATTCATTCATTTCTTTA 992  
539 GAACCTTAAATAGCCTTCAAACTTTTGGCCAGTTCTGTATTCATTCATTTCTTTA 598  
993 CTCTGCTCTCCAGCTT 1009  
599 CTCTGCTCTCCAGCTT 615

RESULT 6  
US-10-027-632-68210  
; Sequence 68210, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006

```

; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68210
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-68210
```

```

Query Match          7.4%; Score 598.2; DB 13; Length 615;
Best Local Similarity 98.7%; Pred. No. 1.2e-133;
Matches 609; Conservative 4; Mismatches 2; Indels 2; Gaps 1;
```

```

QY 393 AACCTGGTGAAGAATTGACACATTTCCCAAAAGTAAAGCATCTGGAGATCCCTG 452
D 1 AACCTGGTGAAGAATTGACACATTTCCCAAAAGTAAAGCATCTGGAGATCCCTG 60
QY 453 TTACCTTCCTGGTATACAGCATCTCCAGCCCCCATATCTTTGCTTTTACTCTAA 512
D 61 TTACCTTCCTGGTATACAGCATCTCCAGCCCCCATATCTTTGCTTTTACTCTAA 120
QY 513 TCATATAGTAAGTCTCATTTGATGCTAGGCCATTTGTAAGCAATTAAGAGAGGA 572
D 121 TCATATAGTAAGTCTCATTTGATGCTAGGCCATTTGTAAGCAATTAAGAGAGGA 180
QY 573 GGCTTCGACACTGAGAGAAATGTCATCTGAAGGTGTCAGACAGCCTGGGCTG 632
D 181 GGCTTCGACACTGAGAGAAATGTCATCTGAAGGTGTCAGACAGCCTGGGCTG 240
QY 633 AGCTTGGCTTACATCTCTCCCAAGTGAAGATCACTGCCCATTTAATCTGTGAA 692
D 241 AGCTTGGCTTACATCTCTCCCAAGTGAAGATCACTGCCCATTTAATCTGTGAA 300
QY 693 CTAAGAGCGAAGCGCTCCCAATGACTTATTTCCCTGCATTGATACCGTCAATCT 752
D 301 CTAAGAGCGAAGCGCTCCCAATGACTTATTTCCCTGCATTGATACCGTCAATCT 360
QY 753 TGAGAAATGTTTCTTTGTTCTCCCTGAGCAAGGTTGAAAAATTTGAAATTTACCTA 812
D 361 TGAGAAATGTTTCTTTGTTCTCCCTGAGCAAGGTTGAAAAATTTGAAATTTACCTA 420
QY 813 GAGACACACATAGTCAATCTGCTGTGTGCTGAATGTCTGCCGCCCGAGTGAAG 872
D 421 GAGACACACATAGTCAATCTGCTGTGTGCTGAATGTCTGCCGCCCGAGTGAAG 480
QY 873 AGTTCTTAAAGCCTATTGTCAACAATACCTTCAGATGTTAGATTTCATTTAAG 932
D 481 AGTTCTTAAAGCCTATTGTCAACAATACCTTCAGATGTTAGATTTCATTTAAG 538
QY 933 GAATTTAAATAGCTTTCAAACTTTTGGCACTTTCTGATATCCATCTATCTTTA 992
D 539 GAATTTAAATAGCTTTCAAACTTTTGGCACTTTCTGATATCCATCTATCTTTA 598
QY 993 CTCTGCTCCCAAGCTT 1009
D 599 CTCTGCTCCCAAGCTT 615
```

```

RESULT 7
US-10-027-632-68211
; Sequence 68211, Application US/10027632
```

```

; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68211
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-68211
```

```

Query Match          7.4%; Score 598.2; DB 13; Length 615;
Best Local Similarity 98.7%; Pred. No. 1.2e-133;
Matches 609; Conservative 4; Mismatches 2; Indels 2; Gaps 1;
```

```

QY 393 AACCTGGTGAAGAATTGACACATTTCCCAAAAGTAAAGCATCTGGAGATCCCTG 452
D 1 AACCTGGTGAAGAATTGACACATTTCCCAAAAGTAAAGCATCTGGAGATCCCTG 60
QY 453 TTACCTTCCTGGTATACAGCATCTCCAGCCCCCATATCTTTGCTTTTACTCTAA 512
D 61 TTACCTTCCTGGTATACAGCATCTCCAGCCCCCATATCTTTGCTTTTACTCTAA 120
QY 513 TCATATAGTAAGTCTCATTTGATGCTAGGCCATTTGTAAGCAATTAAGAGAGGA 572
D 121 TCATATAGTAAGTCTCATTTGATGCTAGGCCATTTGTAAGCAATTAAGAGAGGA 180
QY 573 GGCTTCGACACTGAGAGAAATGTCATCTGAAGGTGTCAGACAGCCTGGGCTG 632
D 181 GGCTTCGACACTGAGAGAAATGTCATCTGAAGGTGTCAGACAGCCTGGGCTG 240
QY 633 AGCTTGGCTTACATCTCTCCCAAGTGAAGATCACTGCCCATTTAATCTGTGAA 692
D 241 AGCTTGGCTTACATCTCTCCCAAGTGAAGATCACTGCCCATTTAATCTGTGAA 300
QY 693 CTAAGAGCGAAGCGCTCCCAATGACTTATTTCCCTGCATTGATACCGTCAATCT 752
D 301 CTAAGAGCGAAGCGCTCCCAATGACTTATTTCCCTGCATTGATACCGTCAATCT 360
QY 753 TGAGAAATGTTTCTTTGTTCTCCCTGAGCAAGGTTGAAAAATTTGAAATTTACCTA 812
D 361 TGAGAAATGTTTCTTTGTTCTCCCTGAGCAAGGTTGAAAAATTTGAAATTTACCTA 420
QY 813 GAGACACACATAGTCAATCTGCTGTGTGCTGAATGTCTGCCGCCCGAGTGAAG 872
D 421 GAGACACACATAGTCAATCTGCTGTGTGCTGAATGTCTGCCGCCCGAGTGAAG 480
QY 873 AGTTCTTAAAGCCTATTGTCAACAATACCTTCAGATGTTAGATTTCATTTAAG 932
D 481 AGTTCTTAAAGCCTATTGTCAACAATACCTTCAGATGTTAGATTTCATTTAAG 538
QY 933 GAATTTAAATAGCTTTCAAACTTTTGGCACTTTCTGATATCCATCTATCTTTA 992
D 539 GAATTTAAATAGCTTTCAAACTTTTGGCACTTTCTGATATCCATCTATCTTTA 598
```

QY	993	CTCTGCGTCCCAAGCTT	1009
Dd	599	CTCTGCCCTCTCAAGCTT	615

```

RESULT 8
US-10-027-633-68742
; Sequence 68742, Application US/10027632
; Publication NO. US2002019837A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIORITY FILING DATE: 2002-04-30
; PRIORITY APPLICATION NUMBER: US 60/219,006
; PRIORITY FILING DATE: 2000-07-12
; PRIORITY APPLICATION NUMBER: US 60/199,676
; PRIORITY FILING DATE: 2000-04-20
; PRIORITY APPLICATION NUMBER: US 60/193,483
; PRIORITY FILING DATE: 2000-03-29
; PRIORITY APPLICATION NUMBER: US 60/185,218
; PRIORITY FILING DATE: 2000-02-24
; PRIORITY APPLICATION NUMBER: US 60/167,363
; PRIORITY FILING DATE: 1999-11-22
; PRIORITY APPLICATION NUMBER: US 60/156,358
; PRIORITY FILING DATE: 1999-09-28
; PRIORITY APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68742
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Human
; US-10-027-633-68742

```

Query Match	7.4%;	Score 598.2;	DB 13;	Length 615;
Best Local Similarity	98.7%;	Pred. No. 1.2e-133;		
Matches 609;	Conservative 4;	Mismatches 2;	Indels 2;	Gaps 1

QY	393	AAGCTTGGTGCAAGAAATTGGACACATTTCCCAAAAGTAAGACATCTGGGAAGTCCCTG	452
Db	1	AAGCTTGGTGCAAGAAATTGGACACATTTCCCAAAAGTAAGACATCTGGGAAGTCCCTG	60
QY	453	TTTACCTTCTGGTATTAAGCATCTCTCAGGCCCATATCTTTGCTTTTAACTCTTAAAA	512
Db	61	TTTACCTTCTGGTATTAAGCATCTCTCAGGCCCATATCTTTGCTTTTAACTCTTAAAA	120
QY	513	TCATTAAGTGAACCTCTCATTAATGCTAGAGCCATTGTAGTAAGAAATTAAGAAAGAGAGGA	572
Db	121	TCATTAAGTGAACCTCTCATTAATGCTAGAGCCATTGTAGTAAGAAATTAAGAAAGAGAGGA	180
QY	573	GGCCTTCTGCAACTGAGAGAAATTTGATCATGTGAAGTGTGCAAGCAGACCTTGAGGCTG	632
Db	181	GGCCTTCTGCAACTGAGAGAAATTTGATCATGTGAAGTGTGCAAGCAGACCTTGAGGCTG	240
QY	633	AACCTTGGCTTACATCTCTGCCCAAGTGAAGATTAAGTGCCTCCATTTTAACTCTGTAGAA	692
Db	241	AACCTTGGCTTACATCTCTGCCCAAGTGAAGATTAAGTGCCTCCATTTTAAATCTGTAGAA	300
QY	693	CTTAAAGAAAGCAAGCCCTGCACACATGACTTAATTTCCCTGCATTTGATACCGTCATCCT	752
Db	301	CTTAAAGAAAGCAAGCCCTGCACACATGACTTAATTTCCCTGCATTTGATACCGTCATCCT	360
QY	753	TGAGAAATGTTTCTTTTGTCTCTCCGTAGCAAAAGTTGGAATAATTTGAAATTTTACCTA	812
Db	361	TGAGAAATGTTTCTTTTGTCTCTCCGTAGCAAAAGTTGGAATAATTTGAAATTTTACCTA	420
QY	813	GAGACCACACATAGTTCATCTCTGCTGTGTGCTGAATGTCTGCCCCCAGTAGAAAC	872

Db	421	GAGACACACATAGTTCACATCTCTGCTGTGTGGCTGAATGTCGCCCCVCAAGTAGGAAC	480
Qy	873	AGTTCCTTTAAAGCCATTTGTGCACAAATACCTTCAGATGTTBACATTTTACAATTTTAAG	932
Db	481	AGTCTCTTCAAGGCCATATGTGCACAAATACCTTCAGATGTTBACATTTT - CATTTAAG	538
Qy	933	GAACTTAAATATGCGCTTCAAACTTTTGGCAGTTTCCTGATATCCAAATCATTTCTTTTA	992
Db	539	GAACTTAAATATGCGCTTCAAACTTTTGGCAGTTTCCTGATATCCAAATCATTTCTTTTA	598
Qy	993	CTCTGCGCTCCCAAGCTT	1009
Db	599	CTCTGCGCTTCAAGCTT	615

```

RESULT 9
US-10-027-632-68743
Sequence 68743, Application US/10027632
Publication No. US20020198372A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
POLYMORPHISMS IN THE HUMAN GENOME
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
PRIOR FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-25
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 68743
LENGTH: 615
TYPE: DNA
ORGANISM: Human
US-10-027-632-68743

```

Query Match	7.4%	Score 558.2;	DB 13;	Length 615;
Best Local Similarity	98.7%;	Pred. No. 1.2e-133;		
Matches 609;	Conservative 4;	Mismatches 2;	Indels 2;	Gaps 1

QY	393	ANGCTTGTCAGAAAGATTGGACACATTTCCCAAAAGTAAGACATCTGGGAAGTCCCTG	452
Db	1	AAGCTTGGTCAGAAAGATTGGACACATTTCCCAAAAGTAAGACATCTGGGAAGTCCCTG	60
QY	453	TTTACCTTCCTGGTTACAGATCCTCCAGGCCCATATCTTGCTTTTGTGCTCTAAAA	512
Db	61	TTTACCTTCCTGGTTACAGATCCTCCAGGCCCATATCTTGCTTTTGTGCTCTAAAA	120
QY	513	TCGAATACTGAATCTCATTTATGTCTAAGCCATTGTAGTAAACATAAAGANGAGGGA	572
Db	121	TGAATPACTGGAATCTCATTTATGTCTAAGCCATTGTAGTAAACATAAAGANGAGGGA	180
QY	573	GGCTTCTGACAACTGAGAGAAATTTGATCTGTAAGTGGTGCAAGACACAGCCTGGAGCTG	632
Db	181	GGCTTCTGACAACTGAGAGAAATTTGATCTGTAAGTGGTGCAAGACACAGCCTGGAGCTG	240
QY	633	AACCTTGGCCATTCCTCTGCCAAGTGAAGATCAGTGGCCCATTTAAACATCTGGTGA	692
Db	241	AACCTTGGCCATTCCTCTGCCAAGTGAAGATCAGTGGCCCATTTAAATCTGGTGA	300
QY	693	CTAAAGAACGACAGCCCTGGCCACATGACTTATTTCCCTGCACTTGTATACCGTCATCCT	752

Db 301 CTAAGAAAGCAACACCGCCGACAAATGACTTATTTCCCTGCAATTTGATACCGTAACTCT 360  
Qy 753 TGAGAAATGTTTCTTTTGTCTCTCCCTAGCAAAAGTTGAAATTTGAAATTTTACCTA 812  
Db 361 TGAGAAATGTTTCTTTTGTCTCTCCCTAGCAAAAGTTGAAATTTGAAATTTTACCTA 420  
Qy 813 GAGACCAACATAGTTACATCTGCTGTGTGCTGAATGTCTGCCCCCAATGAGAAC 872  
Db 421 GAGACCAACATAGTTACATCTGCTGTGTGCTGAATGTCTGCCCCCAATGAGAAC 480  
Qy 873 AGTTCTCTTAAGCCATTTGCAACAATACCTTCAGATGTTTACATTTTAAATTTAAG 932  
Db 481 AGTTCTCTTAAGCCATTTGCAACAATACCTTCAGATGTTTACATTTT--CATTTAAG 538  
Qy 933 GAACCTAAATAGCCTTCAAACTTTTGGCAGTTTCTGTGATTCGAATCTATTTCTTTTA 992  
Db 539 GAACCTAAATAGCCTTCAAACTTTTGGCAGTTTCTGTGATTCGAATCTATTTCTTTTA 598  
Qy 993 CTCTGCTCCCAAGCTT 1009  
Db 599 CTCTGCTCCCAAGCTT 615

## RESULT 10

US-10-027-632-68744  
; Sequence 68744, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 68744  
; LENGTH: 615  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-68744

Query Match 7.4%; Score 598.2; DB 13; Length 615;  
Best Local Similarity 98.7%; Pred. No. 1.2e-133;  
Matches 609; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

Qy 393 AAGCTTGTGCAAAAGATTGGACACTTTCCCAAAAGTAAGCACTAGCGGAAGTCCCTG 452  
Db 1 AAGCTTGTGCAAAAGATTGGACACTTTCCCAAAAGTAAGCACTAGCGGAAGTCCCTG 60  
Qy 453 TTATACCTTCTGTATACAGCACTCTCCAGCCCAATCTTTGCTTTTATCTCTAAAAA 512  
Db 61 TTATACCTTCTGTATACAGCACTCTCCAGCCCAATCTTTGCTTTTATCTCTAAAAA 120  
Qy 513 TCAATACTGAATCTTCATTTGATGTCTAGGCCATTGTATGAACAATAAAGAGAGGA 572  
Db 121 TCAATACTGAATCTTCATTTGATGTCTAGGCCATTGTATGAACAATAAAGAGAGGA 180

Qy 573 GCCTTCTGACACTGAAGGAAATTTGATCTGAAGTGTGCAAGCAAGCCCTGGGGCTG 632  
Db 181 GCCTTCTGACACTGAAGGAAATTTGATCTGAAGTGTGCAAGCAAGCCCTGGGGCTG 240  
Qy 633 AGCCTTGCCCTACATCTCTGCGCAAGTGAAGATCAGTCCCAATTTAATCTGTAGAA 692  
Db 241 AGCCTTGCCCTACATCTCTGCGCAAGTGAAGATCAGTCCCAATTTAATCTGTAGAA 300  
Qy 693 CTAAGAAAGCAAGCCTGCCACATGACTTATTTCCCTGCAATTTATACCGTCACTCT 752  
Db 301 CTAAGAAAGCAAGCCTGCCACATGACTTATTTCCCTGCAATTTATACCGTCACTCT 360  
Qy 753 TGAGAAATGTTTCTTTTGTCTCTCCCTAGCAAAAGTTGAAATTTGAAATTTTACCTA 812  
Db 361 TGAGAAATGTTTCTTTTGTCTCTCCCTAGCAAAAGTTGAAATTTGAAATTTTACCTA 420  
Qy 813 GAGACCAACATAGTTACATCTGCTGTGTGCTGAATGTCTGCCCCCAATGAGAAC 872  
Db 421 GAGACCAACATAGTTACATCTGCTGTGTGCTGAATGTCTGCCCCCAATGAGAAC 480  
Qy 873 AGTTCTCTTAAGCCATTTGCAACAATACCTTCAGATGTTTACATTTTAAATTTAAG 932  
Db 481 AGTTCTCTTAAGCCATTTGCAACAATACCTTCAGATGTTTACATTTT--CATTTAAG 538  
Qy 933 GAACCTAAATAGCCTTCAAACTTTTGGCAGTTTCTGTGATTCGAATCTATTTCTTTTA 992  
Db 539 GAACCTAAATAGCCTTCAAACTTTTGGCAGTTTCTGTGATTCGAATCTATTTCTTTTA 598  
Qy 993 CTCTGCTCCCAAGCTT 1009  
Db 599 CTCTGCTCCCAAGCTT 615

## RESULT 11

US-10-027-632-70068  
; Sequence 70068, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 70068  
; LENGTH: 615  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-70068

Query Match 7.4%; Score 598.2; DB 13; Length 615;  
Best Local Similarity 98.7%; Pred. No. 1.2e-133;  
Matches 609; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

Qy 393 AAGCTTGTGCAAAAGATTGGACACTTTCCCAAAAGTAAGCACTAGCGGAAGTCCCTG 452  
Db 1 AAGCTTGTGCAAAAGATTGGACACTTTCCCAAAAGTAAGCACTAGCGGAAGTCCCTG 60



PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 70070  
LENGTH: 615  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-70070

Query Match  
Best Local Similarity 98.7%; Pred. No. 1,2e-133;  
Matches 609; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

QY 393 AAGCTTGCGCAAGAAATTGACACATTTCCCAAAAGTAAAGACATCTGGGAAGTCCCTG 452  
Db 1 AAGCTTGCGCAAGAAATTGACACATTTCCCAAAAGTAAAGACATCTGGGAAGTCCCTG 60  
QY 453 TTACCTTCCGTATACAGCATCTCCAGCCCATATCTTTTGTAGTCTTAAAAA 512  
Db 61 TTACCTTCCGTATACAGCATCTCCAGCCCATATCTTTTGTAGTCTTAAAAA 120  
QY 513 TCATTAAGTGAATCTCTCATCTGATCTCTAGCCATTGTATTAACAATAAAGAGGGA 572  
Db 121 TCATTAAGTGAATCTCTCATCTGATCTCTAGCCATTGTATTAACAATAAAGAGGGA 180  
QY 573 GGCTTGTGCAACTGAGAGAAATGTGATCTGAAGTGTGCAAGCAAGCTGGGGCTG 632  
Db 181 GGCTTGTGCAACTGAGAGAAATGTGATCTGAAGTGTGCAAGCAAGCTGGGGCTG 240  
QY 633 AGCTTGCGCTACATCTGCGCAAGTGAAGATCACTGCCCCATTAACTCTGTAGAA 652  
Db 241 AGCTTGCGCTACATCTGCGCAAGTGAAGATCACTGCCCCATTAAATCTGTAGAA 300  
QY 693 CTAAGAAGCGAAGCGCTGCCACATGACTTATTTCCCTGATTTGATCCGTCATCT 752  
Db 301 CTAAGAAGCGAAGCGCTGCCACATGACTTATTTCCCTGATTTGATCCGTCATCT 360  
QY 753 TGAGAAATGTTTTCTTTTGTCTCCCTGAGCAAGTTGAAAAATTGAAATTACCTA 812  
Db 361 TGAGAAATGTTTTCTTTTGTCTCCCTGAGCAAGTTGAAAAATTGAAATTACCTA 420  
QY 813 GAGACGACATAGTTTCAATCTGCTGCTGAGTGTGCTGCCCCCGCTGAGAAAC 872  
Db 421 GAGACGACATAGTTTCAATCTGCTGCTGAGTGTGCTGCCCCCGCTGAGAAAC 480  
QY 873 AGTTCTTCTAAGGCTATGTGCAACATACCTTCCAGATGTTAGCATTTTAAAG 932  
Db 481 AGTTCTTCTAAGGCTATGTGCAACATACCTTCCAGATGTTAGCATTTT--CATTTAAG 538  
QY 933 GAACTTAAATAGCTTCAAACTTTTGGCAGTTTCTGATATCCATCTATCTTTTA 992  
Db 539 GAACTTAAATAGCTTCAAACTTTTGGCAGTTTCTGATATCCATCTATCTTTTA 598  
QY 993 CTCTGCTCCCAAGCTT 1009  
Db 599 CTCTGCTCCCAAGCTT 615

## RESULT 14

US-10-027-632-70071  
Sequence 70071, Application US/10027632  
Publication No. US20020198371A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 70071  
LENGTH: 615  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-70071

Query Match  
Best Local Similarity 98.7%; Pred. No. 1,2e-133;  
Matches 609; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

QY 393 AAGCTTGCGCAAGAAATTGACACATTTCCCAAAAGTAAAGACATCTGGGAAGTCCCTG 452  
Db 1 AAGCTTGCGCAAGAAATTGACACATTTCCCAAAAGTAAAGACATCTGGGAAGTCCCTG 60  
QY 453 TTACCTTCCGTATACAGCATCTCCAGCCCATATCTTTTGTAGTCTTAAAAA 512  
Db 61 TTACCTTCCGTATACAGCATCTCCAGCCCATATCTTTTGTAGTCTTAAAAA 120  
QY 513 TCATTAAGTGAATCTCTCATCTGATCTCTAGCCATTGTATTAACAATAAAGAGGGA 572  
Db 121 TCATTAAGTGAATCTCTCATCTGATCTCTAGCCATTGTATTAACAATAAAGAGGGA 180  
QY 573 GGCTTGTGCAACTGAGAGAAATGTGATCTGAAGTGTGCAAGCAAGCTGGGGCTG 632  
Db 181 GGCTTGTGCAACTGAGAGAAATGTGATCTGAAGTGTGCAAGCAAGCTGGGGCTG 240  
QY 633 AGCTTGCGCTACATCTGCGCAAGTGAAGATCACTGCCCCATTAACTCTGTAGAA 692  
Db 241 AGCTTGCGCTACATCTGCGCAAGTGAAGATCACTGCCCCATTAAATCTGTAGAA 300  
QY 693 CTAAGAAGCGAAGCGCTGCCACATGACTTATTTCCCTGATTTGATCCGTCATCT 752  
Db 301 CTAAGAAGCGAAGCGCTGCCACATGACTTATTTCCCTGATTTGATCCGTCATCT 360  
QY 753 TGAGAAATGTTTTCTTTTGTCTCCCTGAGCAAGTTGAAAAATTGAAATTACCTA 812  
Db 361 TGAGAAATGTTTTCTTTTGTCTCCCTGAGCAAGTTGAAAAATTGAAATTACCTA 420  
QY 813 GAGACGACATAGTTTCAATCTGCTGCTGAGTGTGCTGCCCCCGCTGAGAAAC 872  
Db 421 GAGACGACATAGTTTCAATCTGCTGCTGAGTGTGCTGCCCCCGCTGAGAAAC 480  
QY 873 AGTTCTTCTAAGGCTATGTGCAACATACCTTCCAGATGTTAGCATTTTAAAG 932  
Db 481 AGTTCTTCTAAGGCTATGTGCAACATACCTTCCAGATGTTAGCATTTT--CATTTAAG 538  
QY 933 GAACTTAAATAGCTTCAAACTTTTGGCAGTTTCTGATATCCATCTATCTTTTA 992  
Db 539 GAACTTAAATAGCTTCAAACTTTTGGCAGTTTCTGATATCCATCTATCTTTTA 598  
QY 993 CTCTGCTCCCAAGCTT 1009  
Db 599 CTCTGCTCCCAAGCTT 615

## RESULT 15

US-10-027-632-295009  
Sequence 295009, Application US/10027632  
Publication No. US20020198371A1

```

; GENERAL INFORMATION:
; APPLICANT: Mang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 295009
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-295009

```

```

Query Match 7.4% Score 598.2; DB 13; Length 615;
Best Local Similarity 98.7% Pred. No. 1.2e-133;
Matches 609; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

```

```

QY 393 AAGCTTGTCGCAAGAAATTGACACATTTCCCAAAAGTAAGACATCTGGGAAGTCCCTG 452
DB 1 AAGCTTGTCGCAAGAAATTGACACATTTCCCAAAAGTAAGACATCTGGGAAGTCCCTG 60
QY 453 TTTACCTTCTGCTATACAGCATCTCCAGCCCAATCTTGGCTTTTATGTCCTAAAAA 512
DB 61 TTTACCTTCTGCTATACAGCATCTCCAGCCCAATCTTGGCTTTTATGTCCTAAAAA 120
QY 513 TCAATAACTGTAAGTCTCATGATGCTAGAGCATTTAGTAATTAACATTAAGAGAGAGGA 572
DB 121 TCAATAACTGTAAGTCTCATGATGCTAGAGCATTTAGTAATTAACATTAAGAGAGAGGA 180
QY 573 GGCCTTGACAACTGAGAGAAATTGATCTGAAGTGTGCAAGCACAGCTGGGGCTG 632
DB 181 GGCCTTGACAACTGAGAGAAATTGATCTGAAGTGTGCAAGCACAGCTGGGGCTG 240
QY 633 AGCCTTGCCCTACATCTGTCGCCAAGTGAAGATCAGTCCCATTTAATCTGTGTAAG 692
DB 241 AGCCTTGCCCTACATCTGTCGCCAAGTGAAGATCAGTCCCATTTAATCTGTGTAAG 300
QY 693 CTAAGAAGCGAAGCGCTGCACAAATGATTTATTTCCCTGATTTGATACCGTCAATCCT 752
DB 301 CTAAGAAGCGAAGCGCTGCACAAATGATTTATTTCCCTGATTTGATACCGTCAATCCT 360
QY 753 TGAGAAATGTTTCTTTTGTCTCTCCCTGAGCAAGGTTGAAAAATTGAAATTTAAGCTA 812
DB 361 TGAGAAATGTTTCTTTTGTCTCTCCCTGAGCAAGGTTGAAAAATTGAAATTTAAGCTA 420
QY 813 GAGACACACATAGTTCATCTGCTGTGTGCTGAATGCTGCCGCCAGTAGGAAAG 872
DB 421 GAGACACACATAGTTCATCTGCTGTGTGCTGAATGCTGCCGCCAGTAGGAAAG 480
QY 873 AGTTCTTCTAAAGCCTATGTCACAAATACCTTCAGATGTTAGCATTTTACAAATTTAAG 932
DB 481 AGTTCTTCTAAAGCCTATGTCACAAATACCTTCAGATGTTAGCATTTT -CATTTAAG 538
QY 933 GAACCTAAATAGCCTTAAAGCTTTTGGCGAGTTCTGTGATATCCATCTATCTTTTA 992
DB 539 GAACCTAAATAGCCTTAAAGCTTTTGGCGAGTTCTGTGATATCCATCTATCTTTTA 598

```

```

QY 993 CTCTGCTCCCAAGCTT 1009
DB 599 CTCTGCTCTCAAGCTT 615

Search completed: September 12, 2004, 01:42:10
Job time : 2308 secs

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 11, 2004, 10:11:48 ; Search time 12097 Seconds

(without alignments)  
19849.698 Million cell updates/sec

Title: US-09-939-209a-3\_COPY\_4110\_12150

Perfect score: 8041  
1 caaatgaagtcctgggcac.....ctagatctcaggtgaccc 8041

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

```
Database :
1: em_escba:*
2: em_escbm:*
3: em_escin:*
4: em_escmu:*
5: em_escov:*
6: em_escpl:*
7: em_escro:*
8: em_escro:*
9: gb_esc1:*
10: gb_esc2:*
11: gb_esc3:*
12: gb_esc4:*
13: gb_esc5:*
14: gb_esc5:*
15: em_escfun:*
16: em_escfun:*
17: em_escfun:*
18: em_escfun:*
19: em_escfun:*
20: em_escfun:*
21: em_escfun:*
22: em_escfun:*
23: em_escfun:*
24: em_escfun:*
25: em_escfun:*
26: em_escfun:*
27: em_escfun:*
28: gb_esc1:*
29: gb_esc2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	344.8	4.3	578	14	CB294770 12B22014
2	272.8	3.4	298	14	CD511718 AGENCOURT
3	268.4	3.3	864	28	AQ745626 HS 2273 A
4	262.8	3.3	787	12	B1547160 603190405

```

c 5 240.4 3.0 647 29 AG090278
c 6 226.4 2.8 712 28 AQ274674
c 7 194.8 2.4 474 9 AM102832
c 8 189.2 2.4 734 28 AQ282818
c 9 181.2 2.3 852 29 CC904986
c 10 179.2 2.2 437 28 AQ472298
c 11 172.4 2.1 777 12 BG699761
c 12 170.4 2.1 593 12 BG699761
c 13 170.4 2.1 680 12 B1548918
c 14 170.4 2.1 680 12 B1548918
c 15 170.4 2.1 720 12 B1545493
c 16 170.4 2.1 833 12 B1551559
c 17 170.4 2.1 834 12 B1547168
c 18 169.8 2.1 482 28 AQ457908
c 19 169.6 2.1 803 12 B1551703
c 20 168.4 2.1 454 28 AQ466676
c 21 168.4 2.1 727 14 CB398484
c 22 168.4 2.1 902 12 BG708958
c 23 168.2 2.1 902 12 B1550370
c 24 168.2 2.1 174 9 AU076999
c 25 168.2 2.1 568 9 AU280510
c 26 168.2 2.1 574 14 CB155278
c 27 167.4 2.1 958 12 B1552207
c 28 166.8 2.1 618 12 BG708946
c 29 166.8 2.1 778 12 B1666696
c 30 166.8 2.1 845 12 B1668181
c 31 166.2 2.1 521 28 AQ316196
c 32 166.2 2.1 674 9 AU708644
c 33 166.2 2.1 912 12 B1546499
c 34 165.6 2.1 441 28 AQ737144
c 35 165.6 2.1 568 28 AQ757067
c 36 164.2 2.0 549 29 CE636174
c 37 164.2 2.0 712 29 AG181600
c 38 162.8 2.0 490 9 AU024209
c 39 162.6 2.0 643 29 AG162354
c 40 162.4 2.0 391 9 AA933951
c 41 162.2 2.0 446 28 AQ053292
c 42 160 2.0 687 28 AQ211630
c 43 159.8 2.0 618 28 AQ285845
c 44 159.6 2.0 548 28 AQ393760
c 45 159.6 2.0 775 28 AQ895749
```

#### ALIGNMENTS

```

RESULT 1
LOCUS 12B22014 rev.1 B11_r_087.abi 578 bp mRNA linear EST 28-FEB-2003
DEFINITION 12B22014 rev.1 B11_r_087.abi Chimpanzee brain library Koo's Pan
troglodytes cDNA clone 12B22014_rev.1 B11_r_087.abi 5', mRNA
sequence.
ACCESSION CB294770.1 GI:28620200
VERSION CB294770
KEYWORDS EST.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE 1 (bases 1 to 578)
AUTHORS Hellmann, I., Zollner, S., Enard, W., Ebersberger, I., Nickel, B. and
Paabo, S.
TITLE Selection on human genes as revealed by comparisons to chimpanzee
cDNA
JOURNAL Genome Res. (2003) In press
COMMENT Contact: Paabo S
Evolutionary Genetics
Max-Planck-Institute for evolutionary Anthropology
Deutscher Platz 6, 04103 Leipzig, Germany
Tel: +49-(0)-341-3550 500
Fax: +49-(0)-341-3550 555
Email: paabo@eva.mpg.de
Seq primer: M13 reverse.
```

FEATURES  
source

Location/Qualifiers  
1. 578  
/organism="Pan troglodytes"  
/mol\_type="mRNA"  
/db\_xref="taxon:9598"  
/clone="12822014\_rev\_1\_S11\_r\_087.ab1"  
/sex="male"  
/tissue\_type="brain, presumably cortex"  
/dev\_stage="adult"  
/lab\_host="Episcurlan Coli (TM) XL-10-Gold"  
/clone\_lib="Chimpanzee brain library Kooos"  
/note="Vector: pUchi; Site 1: SfiI-A; Site 2: SfiI-B; The library was prepared using the SMART cDNA library construction kit (Clontech), doing only primer extension, but not PCR amplification of the cDNA. The only deviation from the published protocol was that we cloned the cDNA into a plasmid vector."

## ORIGIN

Query Match 4.3%; Score 344.8; DB 14; Length 578;  
Best Local Similarity 99.4%; Pred. No. 5.3e-58;  
Matches 346; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 TTACAGGAGCGAGCATGTGAAGAGAACTGCTAGTGGCTGAATTCGTGTTAC 5665  
Qy 5606 TTACAGGAGCGAGCATGTGAAGAGAACTGCTAGTGGCTGAATTCGTGTTAC 5665  
Qy 5666 CAAGATTTTGAAGAAAGTATTCCTCAGTCAGTTTACAGATATAGCAATCTATT 5725  
Db 61 CAAGATTTTGAAGAAAGTATTCCTCAGTCAGTTTACAGATATAGCAATCTATT 120  
Qy 5726 TCTAGGAGTATTCGTATGCGCGGCTATATACGTCTGCATCAGCTATTCTC 5785  
Db 121 TCTAGGAGTATTCGTATGCGCGGCTATATACGTCTGCATCAGCTATTCTC 180  
Qy 5786 TCCACCTCTTGTGTCATACACCAAGCAACTCCGCAATCAGCTGCGAGAGA 5845  
Db 181 TCCACCTCTTGTGTCATACACCAAGCAACTCCGCAATCAGCTGCGAGAGA 240  
Qy 5846 TGAATCTGCGAGTCCCTTTGGAATGCTGAGATCAGATCTGGACCACTTAAT 5905  
Db 241 TGAATCTGCGAGTCCCTTTGGAATGCTGAGATCAGATCTGGACCACTTAAT 300  
Qy 5906 GATGCTTCTAATCCAAAGAGAAAGAGATTGGAGTCACTCAAG 5953  
Db 301 GATGCTTCTAATCCAAAGAGAGAGATTGGAGTCACTCAAG 348

RESULT 2 298 bp mRNA linear EST 06-JUN-2003  
CD511718/c AGENCOURT 14360735 NIH MGC 187 Homo sapiens cDNA clone  
LOCUS IMAGE:30405362 5', mRNA sequence.  
DEFINITION CD511718  
ACCESSION CD511718.1 GI:31443436  
VERSION  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgs.nci.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Daniela S. Gerhard, Ph.D.  
COMMENT Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgaabs-r@mail.nih.gov

Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation

FEATURES  
source

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: NDCM132 row: 9 column: 03  
High quality sequence stop: 298.  
Location/Qualifiers  
1. 298  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30405362"  
/lab\_host="DH10B (TI page-resistant)"  
/clone\_lib="NIH-MGC\_187"  
/note="Organ: Blood vessels - aorta, basilar and artery; Vector: pDNR-LIB; Site 1: SfiI (ggccatcattggcc); Site 2: SfiI (ggccgcctcgcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence:  
5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:  
5'-ATTCTAGAGCGCCGAGCGCGGAGATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA).  
Note: this is a NIH-MGC Library."

## ORIGIN

Query Match 3.4%; Score 272.8; DB 14; Length 298;  
Best Local Similarity 97.5%; Pred. No. 1.2e-43;  
Matches 277; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2791 TTTTCTTAAATTTTAAATTTGTATATATTTATGGGTATATAGGAATCTTTTGA 2850  
Db 286 TTTTCTTAAATTTTAAATTTGTATATATTTATGGGTATATAGGAATCTTTTGA 227  
Qy 2851 TGCATAGGTTGATATGATATATATATATGAGGCTTTTAAAGTATTCATCAGTATGAT 2910  
Db 226 TGCATAGGTTGATATGATATATATATGAGGCTTTTAAAGTATTCATCAGTATGAT 167  
Qy 2911 ACATTGACCCCTTAATATTTCTCAGATCCGCTGATCTTGGCCCTGGATTTGA 2970  
Db 166 ACATTGACCCCTTAATATTTCTCAGATCCGCTGATCTTGGCCCTGGATTTGA 107  
Qy 2971 TCACCTGATGATGATGATGATGATCCCTTAATATTTCTCAGATCCGCTGATCTT 3030  
Db 106 TCACCTGATGATGATGATGATGATCCCTTAATATTTCTCAGATCCGCTGATCTT 47  
Qy 3031 CCCCTGGGATTTGATCAGCTGATGATGATGATGATGATGATGATGATGATGAT 3074  
Db 46 CCCCTGGGATTTGATCAGCTGATGATGATGATGATGATGATGATGATGATGAT 3

RESULT 3 864 bp DNA linear GSS 16-JUL-1999  
A0745626/c HS\_2273\_A2\_B06\_T7C CIT Approved Human Genomic Sperm Library D Homo  
LOCUS sapiens genomic clone Plate=2273 Col=12 Row=C, genomic survey  
DEFINITION A0745626  
ACCESSION A0745626  
VERSION A0745626.1 GI:5523148  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
TITLE Mahairas G.G., Wallace J.C., Smith K., Swartzell S., Holzman T.,  
Keller A., Shaker R., Furlong V., Young D., Zhao S., Adams M.D. and  
Hood L.  
Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
JOURNAL MEDLINE 99380589  
PUBMED 10449764

## COMMENT

Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones may be purchased from Research Genetics (info@resgen.com).  
BAC end Web Server: <http://www.htcsc.washington.edu>  
Plate: 2273 Row: C Column: 12  
Seq primer: T7  
Class: BAC ends  
High quality sequence stop: 864.

## FEATURES

source

1. 864  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="Plate=2273 Col=12 Row=C"  
/sex="male"  
/clone\_lib="CIT Approved Human Genomic Sperm Library D"  
/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"

## ORIGIN

Query Match 3.3%; Score 268.4; DB 28; Length 864;

Best Local Similarity 70.7%; Pred. No. 8e-43;

Matches 454; Conservative 0; Mismatches 156; Indels 32; Gaps 6;

1139 TTCATTATCTCTCTCTCTCTCTCAAGCTGACGAGGCTGAAAGCTGTATATTT 1198  
703 TTTAAATATGTGTTGTAATTCCTATTAAGGCTGAAAGCTGTATATCTTT 644  
1139 TCTCTCATCATTAAGGCTGCAACCAAACTCTTATAGTAAAGACAGTTATAGAG 1258  
643 CCTCACCCCTTATATCATCATCAAGCAACTCTTATTAACAAACAGATTAACAGAA 584  
1259 CAAAACCTAACAAATTTATTAATCAAGTTTACATGACATGGAGCTTCAAGAAATGA 1318  
583 AAA-----AACATATTTCTTATATCAAACTTTACATGACATGAGGCTTCAGAAATGA 529  
1319 AGACCCAAAGCCCGGGAAGACTGCTGTTTTCGAGGCTGAGATGAGAAATGGA 1378  
528 AAACCTAAATCCCGGGAAGAAATATTTG-TTTTATATCTTAGGTTTAAAGAAATAGA 470  
1379 TAGCATGTAGCCATGTAGATTAGACAAAGAGATAAGTATAGTGAAGACTCAAGG- 1437  
469 CAGCTGTAGAAATGTATGATTAGCAAAAGGTAATCTATGATGTAAGACTGAGGGA 410  
1438 GGAACACAGCAAGGCTGTCTATTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1497  
409 GGAGATCCAGCAAAAGCTGTCTGTTCAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 359  
1498 CATCT 1557  
358 ACATCT 299  
1558 GGGAGAA-----AGTGGCTTTTATGATTTTATGCTTCTCTCTCTCTCTCTCTCTCTCT 1610  
298 TGGAAAGGGAAGAGAGTGTAGCTTTCTAGATTTGATGCTTCTCTCTCTCTCTCTCTCTCT 239  
1611 TGGTTCTATGACCATCTTGGGGAAGGAAATCTGTTCTGATGCTCTCTCTCTCTCTCTCT 1670  
238 TATCTTTCTATGACCATTTAGGGAATGGAATTTGCTCTCTCTCTCTCTCTCTCTCTCTCT 179  
1671 AGAAGAGAGATTAAGGAGAGAGGAGAGAGATGTCAGAAAGACT-----TG 1721  
178 AGAAGAGAGGAG 119  
1722 GGTGCTTCTGAGAGGCT 1763  
118 GGTGTTTCTGAGAGCT 77

## RESULT 4

BIS47160

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BIS47160 787 bp mRNA linear EST 05-SEP-2001  
603190405F1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:5261664 5',  
RNA sequence.  
BIS47160  
BIS47160.1 GI:15434472  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 787)  
NIH-MGC <http://mgs.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroaki  
Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LHAM1659 row: j column: 01  
High quality sequence stop: 721.

## FEATURES

source

1. 787  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5261664"  
/tissue\_type="hippocampus"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_95"  
/note="Organ: brain; Vector: pBluescript (modified  
pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI  
(gtagag); Oligo-dT primed using primer  
5'-TTTTTTTTTTTTTTN-3', size-selected for average  
insert size 2.5 kb and normalized to 10^5. This is a  
primary library enriched for full-length clones and  
constructed using the Cap-trapper method (Carninci, in  
preparation). Library constructed by M. Brownstein  
(NIH/NHGRI, National Institutes of Health). Note: this  
is a NIH\_MGC Library."

## ORIGIN

Query Match 3.3%; Score 262.8; DB 12; Length 787;

Best Local Similarity 99.2%; Pred. No. 1.1e-41;

Matches 264; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

6223 GCAGAGGAGACAGAGAGAGCTGTACTGCAAGAGGCTGTGATTTGGCTGACGCTCGT 6282  
4 GGAAGGGGAGACAGAGAGAGCTGTACTGCAAGAGGCTGTGATTTGGCTGACGCTCGT 63  
6283 AGCTGGGCTATTAAGAAGACCCCTTACAGGCTTTAGCAGAGAGAGCTGAGAGATTGAC 6342  
64 AGCTGGGCTATTAAGAAGACCCCTTACAGGCTTTAGCAGAGAGAGCTGAGAGATTGAC 123  
6343 AATATCTTACCGAGAGAGAGCAAGTACGCTCAAGCCGAGACCAAGCTCTCTCTG 6402  
124 AATATCTTACCGAGAGAGAGCAAGTACGCTCAAGCCGAGACCAAGCTCTCTCTG 183  
6403 CGCATTTCTTCTCTGCTTGGCAATTCAGGCTGTAATTAATGAATGCAAGAGGCTTGA 6462  
184 CGCATTTCTTCTCTGCTTGGCAATTCAGGCTGTAATTAATGAATGCAAGAGGCTTGA 243  
6463 GGTGCGGCGCTCTGCTTGAAGAG 6488  
244 GGTGCGGCGCTCTGCTTGAAGAG 269

```
RESULT 5
AG090278/c 647 bp DNA linear GSS 03-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-089M15.R, genomic survey sequence.
ACCESSION AG090278
VERSION AG090278.1 GI:16642080
KEYWORDS GSS
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.
REFERENCE
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoxi,Y., Watanabe,H. and Sakaki,Y.
TITLE BAC end sequences of Library PTB
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 647)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoxi,Y., Watanabe,H. and Sakaki,Y.
TITLE Submitted (02-NOV-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suohiro-cho, Tsukuba, Ibaraki, Japan
DEFINITION 15-mbp BAC library, Genomic Sciences Center (GSC),
1-7-22 Suohiro-cho, Tsukuba, Ibaraki, Japan
ACCESSION 15-mbp BAC library, Genomic Sciences Center (GSC),
1-7-22 Suohiro-cho, Tsukuba, Ibaraki, Japan
VERSION 15-mbp BAC library, Genomic Sciences Center (GSC),
1-7-22 Suohiro-cho, Tsukuba, Ibaraki, Japan
KEYWORDS BAC library, Genomic Sciences Center (GSC),
1-7-22 Suohiro-cho, Tsukuba, Ibaraki, Japan
SOURCE Clones are derived from the chimpanzee BAC library PTB. This BAC end
clone tracking errors.
PRIMERS
Sequencing: M13rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1..647
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-089M15.R"
/sex="male"
/cell_type="lymphoblast"
/clone_id="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match 3.0%; Score 240.4; DB 29; Length 647;
Best Local Similarity 69.3%; Pred. No. 3.1e-37;
Matches 424; Conservative 0; Mismatches 161; Indels 27; Gaps 6;
QY 1194 CATTTCCTTCATATAGGGTCGACCAAACTCCTTACTTAAAGACAGTTAAT 1253
DB 647 CTTTCCTCCTACCATATAGGGTCGACCAAACTCCTTACTTAAAGACAGTTAAT 588
QY 1254 AAGAGCAAACTTAAATTTATTTAATCAAAAGTTTACATGACATGAGGCTTTAGA 1113
DB 587 CAGAGAAAGCATATATTTATTTAATGAGTTGACATGACATGAGGCTTTAGG 528
QY 1314 ATGAAGACCCAAAGACCCAGGGGAACTGCTGTTTTTTTGTGAGGTTGATGAGA 1373
DB 527 ATGAAGACCCAAAGACCCAGGGGAACTGCTGTTTTTTTGTGAGGTTGATGAGA 476
QY 1374 ATGATAGCATGAGCATGATGATGATGACAAAGATATGATGATGATGATGATGAT 1433
DB 475 ACAGAGGCGCATATGAAATGCGCCAAAGATGCTGTTCTAAGGTAAGACCGAGTG 416
QY 1434 AGGGGAAACACAGAGAGCCCTGCTATTCAGATTTCTTGTGATCTCTCTCTATGT 1493
DB 415 GAGAAGGCGAGCGGAGAGAGCTGTGTTTCAGATTTCTTGTGATCTCTCTCTATGT 364
QY 1494 ATAGCATTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1553
|||||
```

```
DB 363 GTAGCATTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 304
QY 1554 AGAAGGAGAAAGTGGCCCTTTTATGATTTATGCTCTGTTGGGAGAGAGCTTTAG 1613
DB 303 AGAAGGAGAGAGAGATGATTTTGTATGCTTATGCTGTTGGGAGAGAGAGCTTTAG 245
QY 1614 TTTCTATGACCATTTTGGGAGAGAGAGATTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1673
DB 244 TTTCTATGACCATTTTGGGAGAGAGAGATTTCTGCTGCTGCTGCTGCTGCTGCTGCT 185
QY 1674 AAGAGAGTAAAGGAGAGAGAGAGAGAGAGATGTCAGAAAAGACTTGCTGCTGAG 1733
DB 184 AAGAGAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 132
QY 1734 GGCCTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1793
DB 131 ATCTTCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 75
QY 1794 ATATGCTTTCTT 1805
DB 74 GTATGCTCTCT 63
RESULT 6
AQ274674 712 bp DNA linear GSS 26-JAN-2001
LOCUS AQ274674
DEFINITION RPCI-6-112A10Sp6 RPCI-6 Homo sapiens genomic clone
ACCESSION AQ274674
VERSION AQ274674.1 GI:3851781
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE BAC and PAC end sequence database for sequence-ready map building
JOURNAL Unpublished (1998)
AUTHORS Contact: de Jong, P.J.
Children's Hospital Oakland Research Institute
747 Fifty Second Street, Oakland, CA 94609-1809, USA
Tel: 510 450 7911
Fax: 510 450 7924
Email: pdejong@mail.cho.org
PAC end sequences. For clone availability please contact Pieter de
Jong (pdejong@mail.cho.org). BACPAC Resources WWW site:
www.choi.org/bacpac
Seq primer: Sp6
Class: PAC end.
FEATURES
source
1..712
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RPCI-6-112A10Sp6"
/sex="female"
/cell_type="lymphocytes"
/lab_host="E. coli DH10B"
/clone_id="RPCI-6"
/note="Vector: pPAC4; PAC clones in E. coli DH10B. For PAC
library availability, please contact Pieter de Jong
(BACPACResources@mail.cho.org). Clones may be purchased from
BACPAC Resources (www.choi.org/bacpac) or from Research
Genetics (info@resgen.com)"
ORIGIN
Query Match 2.8%; Score 226.4; DB 28; Length 712;
Best Local Similarity 71.5%; Pred. No. 1.9e-34;
Matches 418; Conservative 0; Mismatches 137; Indels 30; Gaps 8;
QY 176 GGCCTGAAAGGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1235
|||||
```

[illegible]

```

source
1. 474
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2596085"
/tissue_type="tumor; 5 pooled (see description)"
/lab_host="DH10B"
/clone_lib="NCI CGAP v23"
/note="Organ: Ovary; Vector: PCMV-SPORT6; Site: 1: SalI; Site: 2: NotI; Cloned unidirectionally. Primer: oligo dT. Average insert size 1.35 kb. Tumor types include: mixed Mullerian tumor, papillary serous, clear cell, spindle cell. All are primary tumors, metastasis positive. Life Technologies catalog # 11534-013"

ORIGIN
Query Match      2.4%; Score 194.8; DB: 9; Length 474;
Best Local Similarity 73.5%; Pred. No. 3,8e+28;
Matches 322; Conservative 0; Mismatches 97; Indels 19; Gaps 5;

Oy      1271 AATTATTTAATCAAGTTTACATCATGACATGGAGCTTCAGAAATGAAGCCCAAAGC 1330
Db      19  AACAAATTTATTTAAGTTTACTATTGCATAGAG-CTTCAGAAATGAAGACC-AAAGA 75

Oy      1331 CCAGGGGAACGTCTGTTTTTTTTTGCTGAGTTCCATGAAGATGATGATGATGCC 1390
Db      76  CCCAGGAAAAATGATCATATTGTATGCTTATGTTATGAAGATGACAGCCATATAGA 135

Oy      1391 AGGTGATTAGCAAAAAGATATGATATGATGATAAGACT-CAGGGGGAACAACAGCA 1449
Db      136  AATGTTTGACAAAAGGATGACATTTGATGATGATGATGATGATGACATGCGAGGAAAACCCAGCA 195

Oy      1450 AGGCGTGTCTATTCAGATTCTTCTTGATCTCTCTCTCTATGATATGATTCATTTCTC 1509
Db      196  AGGCGTGTCTTTCAGATCTTCTTCTGAGCCCTC-----TGTGATGATTCATTTCTC 247

Oy      1510 CTGAGTAGGGGAGAGACTCTCTCTCAATAGAGGCTCTCAAGGAG-----AAGGGAG 1562
Db      248  CTGGGTTGGGACAGAGACCCCTTGCAATAGAGGTCTTCAAGAGAGAGGGCAAGGGAG 307

Oy      1563 AAAGTGCCCTTTTGATTTATGCTGCTTGGGAGAGAGAGAGTCTAGTTTCTATGA 1622
Db      308  AGATGACCTTTCTAGTGTTTTGGTCTTCTTGGGGGATAGGGGTTCTAGTTCTATGG 367

Oy      1623 CCCATTTGGGGAGAGAGAAATTCGTGTTTCTGTGACTTGTCTTTCATGAAGAAAGAGACT 1682
Db      368  CCGGCTTGGGAGAGAAATTTGGTTTCTATGCTCTCACTTCAGGGGAGAAAGAGAGT 427

Oy      1683 AAGAGCAGAGAGGGCAG 1700
Db      428  GAGAGACAGAGAGGGCGG 445

RESULT 8
AQ282818          734 bp   DNA       linear    GSS 27-APR-1999
LOCUS             RPC111-89G8.TV RPC1-11 Homo sapiens genomic clone RPC1-11-89G8,
DEFINITION
ACCESSION         AQ282818
KEYWORDS          Genomic survey sequence.
VERSION           AQ282818.1 GI:3909290
SOURCE            GSS.
ORGANISM          Homo sapiens (human)
REFERENCE         1. Adams,M.D., Rounsley,S.D., Zhao,S., Baas,S., Linher,K., Golden,K.,
AUTHORS           Berry,K., Granger,D., Sub,E., Wible,C., de Jong,P. and Venter,J.C.
TITLE             Use of human BAC End Sequences for Sequence-Ready Map Building
JOURNAL           Unpublished (1998)
COMMENT           Other GSSs: RPC11-89G8.TU
CONTACT           Mark Adams
DEPARTMENT        Eukaryotic Genomics
```



Db	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Oy	3868	AGTGAAGAAAGTAACTTAACTCATGTGAAC	3897									
Db	102	AGGTTGCATGCCTGCAGGTTGATGTGACG	73									
	RESULT 10											
	LOCUS	AO472918	437 bp	DNA	linear	GSS	23-APR-1999					
	DEFINITION	CITB1-El-258994.TR	CITB1-El	Homo sapiens	genomic clone	258994,						
	ACCESSION	AO472918										
	VERSION	AO472918.1	GI:4656484									
	KEYWORDS	GSS.										
	SOURCE	Homo sapiens	(human)									
	ORGANISM	Homo sapiens										
	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.										
	AUTHORS	1 (bases 1 to 437)										
		Zhao,S., Adams,M.D., Niernan,W., Malek,J., Shizuya,H., Simon,M. and Venter,J.C.										
	TITLE	Use of BAC End Sequences from CalTech Libraries for Sequence-Ready										
	JOURNAL	Map Building										
	COMMENT	Unpublished (1997)										
		Other GSSs: CITB1-El-258994.TF										
		Contact: Shaying Zhao, William Niernan, Mark Adams										
		Department of Eukaryotic Genomics										
		The Institute for Genomic Research										
		9712 Medical Center Dr., Rockville, MD 20850										
		Tel: 301 838 0200										
		Fax: 301 838 0208										
		Email: hbae@igr.org										
		end search page:										
		http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html.										
		Seq primer: M13 Reverse										
		Class: BAC ends.										
	FEATURES											
	source	Location/Qualifiers										
		1..437										
		/organism="Homo sapiens"										
		/mol_type="genomic DNA"										
		/db_xref="taxon:9606"										
		/clone="258994"										
		/sex="male"										
		/cell_type="sperm"										
		/clone_1b="CITB1-El"										
		/note="Vector: pBelorAC11; Site_1: EcoRI; Site_2: EcoRI;										
		Caltech Human BAC Library D"										
	ORIGIN											
	Query March	2.2%;	Score 179.2;	DB 28;	Length 437;							
	Best Local Similarity	72.9%;	Pred. No. 4.9e-25;									
	Matches 334;	Conservative	0;	Mismatches 98;	Indels 26;	Gaps 7;						
Oy	1186	GTGTGATTAACATTTTCTTCATCATTAAGGGTGGCAACCAAACTCCTATAGTAAAGAC	1245									
Db	436	GTGTGATTAACATTTTCTTCATCATTAAGGGTGGCA										

QY	1482	CTCTCTCTATGATATGCAATCTTTTCCCTCCCGAGATATAGGGGGGAGGACTCTCTCTCAATAG	1541
DB	149	C-----TGCATATGATATTTATTTCTCTGGGTATAGGGGCGGACTCTCTTAAATGGAG	98
QY	1542	GGCTTTCAGAGGAGAGGAGGAGAAAGTGGCCCTTTTATGATTTATGGCTTCCTTCGGGAGA	1601
DB	97	GGGCTTCAAGGAGAGAGAGAGAGAGAGTACCTTTCTTAGGTTTAT-GTTGCTTTGAGGGA	39
QY	1602	GAGGAGCTTCTAGTTCTTCTATGACCATTTTGGGGGAAG	1639
DB	38	GAGGAGCTCTAGTTTCTTAAAGCTGGCTTTGAGGAAG	1
RESULT 11			
EG699761		777 bp	mRNA linear EST 07-MAY-2001
EG699761			
LOCUS			
DEFINITION	602681454F1 NIH_MGC_95 Homo sapiens CDNA clone IMAGE:4814340 5',		
ACCESSION	EG699761		
VERSION	EG699761.1	GI:13968395	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Euteheria; Primates; Carnivora; Catarrhini; Hominoidea; Homo.		
TITLE	1 (bases 1 to 777)		
JOURNAL	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strassburg, Ph.D.		
	Email: cga@bcr-remail.nih.gov		
	Tissue Procurement: Miklos Palokovits, M.D., Ph.D.		
	CDNA library Preparation: Michael J. Brownstein (NHGRI), Shiroki		
	Toshiyuki and Piero Carninci (RIKEN)		
	CDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	<a href="http://image.llnl.gov">http://image.llnl.gov</a>		
	plate: LLM10709 row: k column: 13		
FEATURES			
source	location/Qualifiers		
	1..777		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4814340"		
	/tissue_type="hippocampus"		
	/lab_host="RDH10B"		
	/clone_id="NIH_MGC_95"		
	/note="Organ: brain; Vector: pBluescriptR (modified		
	pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI		
	(gcgcag); Oligo-dT primed using primer		
	5'-TTTTTTTTTTTTTTTAA-3', size-selected for average		
	insert size 2.5 kb and normalized to R0.5. This is a		
	primary library enriched for full-length clones and		
	constructed using the Cap-trapper method (Carninci, in		
	preparation). Library constructed by M. Brownstein		
	(NIH/NHGRI, National Institutes of Health). Note: this		
	is a NIH_MGC Library."		
ORIGIN			
Query Match	2.1%; Score 172.4; DB 12; Length 777;		
Best Local Similarity	99.4%; Pred. No. 1e-23; 1; Indels 0; Gaps 0;		
Matches 173; Conservative	0; Mismatches		
QY	6315	AGCAGGAGACCCCTAGAGGATTCGACAAATATCTTTACGAGAAAGGCAAGTACGC	6374
DB	1	AGCGGAGAAAGCCTCAGAGGATTCGACAAATATCTTTACGAGAAAGGCAAGTACGC	60
QY	6375	TCAAACCCGAAAGCCAGCTCTCTCTGCGGATTTCTTTCTGCTTGGCAATTCAGAGT	6433
DB	61	TCAAACCCGAAAGCCAGCTCTCTCTGCGGATTTCTTTCTGCTTGGCAATTCAGAGT	120

QY 6435 GTTAATAATGATGTGCAAGGCTTGACAGTCTGCCGGCTTCTTGCTTGAGGAG 6488  
 |||||||  
 Db 121 GTTAATAATGATGTGCAAGGCTTGACAGTCTGCCGGCTTCTTGCTTGAGGAG 174

RESULT 12  
 BG699541 593 bp mRNA linear EST 07-MAY-2001  
 LOCUS 602679246F1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:4812059 5',  
 DEFINITION mRNA sequence.

ACCESSION BG699541  
 VERSION BG699541.1 GI:13967945  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgs.nci.nih.gov/  
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: csapbs-r@mail.nih.gov  
 Tissue Procurement: Miklos Pal Kovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
 Toshiyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LLAM10703 row: 1 column: 12  
 High quality sequence stop: 591.

## FEATURES

Location/Qualifiers  
 1..593  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4812059"  
 /tissue\_type="hippocampus"  
 /lab\_host="DH10B"  
 /clone\_id="NIH\_MGC\_95"  
 /note="Organ: brain; Vector: pBluescript (modified  
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI  
 (gtcgag); Oligo-dT primed using primer  
 5'-TTTTTTTTTTTVA-3', size-selected for average  
 insert size 2.5 kb and normalized to ROP 5. This is a  
 primary library enriched for full-length clones and  
 constructed using the Cap-trapper method (Carninci, in  
 preparation). Library constructed by M. Brownstein  
 (NIH/NHGRI, National Institutes of Health). Note: this  
 is a NIH\_MGC Library."

## ORIGIN

Query Match 2.1%; Score 170; DB 12; Length 593;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-23;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6319 GGAAGACGCTCAGAGATCTGACATATCTTACCGGAGAGGCAAGTACGCTCAA 6378  
 |||||||  
 Db 4 GGAAGACGCTCAGAGATCTGACATATCTTACCGGAGAGGCAAGTACGCTCAA 63  
 |||||||  
 QY 6379 AGCCGAAGCCACAGCTCTCTCGCCGCAATTCCTGCTTGGAATTCAGCTGTTA 6438  
 |||||||  
 Db 64 AGCCGAAGCCACAGCTCTCTCGCCGCAATTCCTGCTTGGAATTCAGCTGTTA 123  
 |||||||  
 QY 6439 AATAAGATGTGCAAGGCTTGACAGTCTGCCGGCTTCTTGCTTGAGGAG 6488  
 |||||||  
 Db 124 AATAAGATGTGCAAGGCTTGACAGTCTGCCGGCTTCTTGCTTGAGGAG 173  
 |||||||

RESULT 13

B1548918 624 bp mRNA linear EST 05-SEP-2001  
 LOCUS 603189043F1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:5260403 5',  
 DEFINITION mRNA sequence.

ACCESSION B1548918  
 VERSION B1548918.1 GI:15436217  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgs.nci.nih.gov/  
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: csapbs-r@mail.nih.gov  
 Tissue Procurement: Miklos Pal Kovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
 Toshiyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LLAM11656 row: e column: 12  
 High quality sequence stop: 624.

## FEATURES

Location/Qualifiers  
 1..624  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5260403"  
 /tissue\_type="hippocampus"  
 /lab\_host="DH10B"  
 /clone\_id="NIH\_MGC\_95"  
 /note="Organ: brain; Vector: pBluescript (modified  
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI  
 (gtcgag); Oligo-dT primed using primer  
 5'-TTTTTTTTTTTVA-3', size-selected for average  
 insert size 2.5 kb and normalized to ROP 5. This is a  
 primary library enriched for full-length clones and  
 constructed using the Cap-trapper method (Carninci, in  
 preparation). Library constructed by M. Brownstein  
 (NIH/NHGRI, National Institutes of Health). Note: this  
 is a NIH\_MGC Library."

## ORIGIN

Query Match 2.1%; Score 170; DB 12; Length 624;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-23;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6319 GGAAGACGCTCAGAGATCTGACATATCTTACCGGAGAGGCAAGTACGCTCAA 6378  
 |||||||  
 Db 4 GGAAGACGCTCAGAGATCTGACATATCTTACCGGAGAGGCAAGTACGCTCAA 63  
 |||||||  
 QY 6379 AGCCGAAGCCACAGCTCTCTCGCCGCAATTCCTGCTTGGAATTCAGCTGTTA 6438  
 |||||||  
 Db 64 AGCCGAAGCCACAGCTCTCTCGCCGCAATTCCTGCTTGGAATTCAGCTGTTA 123  
 |||||||  
 QY 6439 AATAAGATGTGCAAGGCTTGACAGTCTGCCGGCTTCTTGCTTGAGGAG 6488  
 |||||||  
 Db 124 AATAAGATGTGCAAGGCTTGACAGTCTGCCGGCTTCTTGCTTGAGGAG 173  
 |||||||

RESULT 14  
 B1548997 680 bp mRNA linear EST 05-SEP-2001  
 LOCUS 603188958F1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:5260358 5',  
 DEFINITION mRNA sequence.  
 ACCESSION B1548997  
 VERSION B1548997.1 GI:15436309  
 KEYWORDS EST.



SOURCE  
ORGANISM Homo sapiens (human)  
COMMENT  
REFERENCE 1 (bases 1 to 680)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>  
Plate: L1M11656 row: c column: 15  
High quality sequence stop: 680.  
Location/Qualifiers  
1. 680  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5260358"  
/tissue\_type="hippocampus"  
/lab\_host="DH10B"  
/clone\_id="NIH\_MGC\_95"  
/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtagag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.5 kb and normalized to R0T 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH\_MGC Library."  
ORIGIN  
Query Match 2.1%; Score 170; DB 12; Length 680;  
Best Local Similarity 100.0%; Pred. No. 3.1e-23;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6319 GGAAGACGCTCAGAGATTCTGACATATCTTTACCGGAGAGGCAAGTACGCTCAA 6378  
DB 4 GGAAGACGCTCAGAGATTCTGACATATCTTTACCGGAGAGGCAAGTACGCTCAA 63  
QY 6379 AGCCGAAGCCACAGCTCTCTCGCGCATTTCTTCTGCTTGGCAATTCAGCTGTTA 6438  
DB 64 AGCCGAAGCCACAGCTCTCTCGCGCATTTCTTCTGCTTGGCAATTCAGCTGTTA 123  
QY 6439 AATAAGATGTGCAAGGCTTGACAGTCTGCCGCTTCTTGCTTGAGAG 6488  
DB 124 AATAAGATGTGCAAGGCTTGACAGTCTGCCGCTTCTTGCTTGAGAG 173  
RESULT 15  
LOCUS B1545493 720 bp mRNA linear EST 05-SEP-2001  
DEFINITION 603187655F1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:5259084 5',  
B1545493  
B1545493  
B1545493.1 GI:15432805  
ACCESSION B1545493  
VERSION B1545493  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
COMMENT Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 720)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>  
Plate: L1M11652 row: n column: 13  
High quality sequence stop: 714.  
Location/Qualifiers  
1. 720  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5259084"  
/tissue\_type="hippocampus"  
/lab\_host="DH10B"  
/clone\_id="NIH\_MGC\_95"  
/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtagag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.5 kb and normalized to R0T 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH\_MGC Library."  
ORIGIN  
Query Match 2.1%; Score 170; DB 12; Length 720;  
Best Local Similarity 100.0%; Pred. No. 3.1e-23;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6319 GGAAGACGCTCAGAGATTCTGACATATCTTTACCGGAGAGGCAAGTACGCTCAA 6378  
DB 4 GGAAGACGCTCAGAGATTCTGACATATCTTTACCGGAGAGGCAAGTACGCTCAA 63  
QY 6379 AGCCGAAGCCACAGCTCTCTCGCGCATTTCTTCTGCTTGGCAATTCAGCTGTTA 6438  
DB 64 AGCCGAAGCCACAGCTCTCTCGCGCATTTCTTCTGCTTGGCAATTCAGCTGTTA 123  
QY 6439 AATAAGATGTGCAAGGCTTGACAGTCTGCCGCTTCTTGCTTGAGAG 6488  
DB 124 AATAAGATGTGCAAGGCTTGACAGTCTGCCGCTTCTTGCTTGAGAG 173  
Search completed: September 11, 2004, 20:08:10  
Job time : 12103 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 11, 2004, 09:15:38 / Search time 19840 Seconds

(without alignments)  
17566.600 Million cell updates/sec

Title: US-09-939-209A-3\_COPY\_4110\_12150

Perfect score: 8041

Sequence: 1 caaatgaagctctgggac.....cttagatccagctggatcc 8041

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 3470272 segs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_scs:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_inv:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_pln:\*  
35: em\_hcg\_rtd:\*  
36: em\_hcg\_mam:\*  
37: em\_hcg\_vrt:\*  
38: em\_sy:\*  
39: em\_hcgo\_hum:\*  
40: em\_hcgo\_mus:\*  
41: em\_hcgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	* Match	Length	DB	ID	Description
1	8041	100.0	20300	6	AX451337	AX451337 Sequence
2	8039.4	100.0	165329	9	AL583850	Human DNA
3	8009.2	99.6	191699	2	AC031977	AC031977 Homo sapi
4	426	5.3	275631	2	AC125563	AC125563 Rattus no
5	402.6	5.0	224486	2	AC115766	AC115766 Mus muscu
6	326.2	4.1	172931	9	AC006581	AC006581 Homo sapi
7	317	3.9	94158	9	AC119619	AC119619 Homo sapi
8	317	3.9	175440	2	AC137626	AC137626 Homo sapi
9	311.2	3.9	154125	2	AC073227	AC073227 Homo sapi
10	311.2	3.9	196832	9	AC018694	AC018694 Homo sapi
11	307.2	3.8	176773	2	AC009857	AC009857 Homo sapi
12	305.6	3.8	89000	9	AP003059	AP003059 Homo sapi
13	299.4	3.7	159972	9	AL533284	AL533284 Human DNA
14	297.6	3.7	121720	9	AL591491	AL591491 Human DNA
15	293.8	3.7	14937	9	AC010244	AC010244 Homo sapi
16	293.8	3.7	170219	9	AC091991	AC091991 Homo sapi
17	293.8	3.7	182049	2	AC021239	AC021239 Homo sapi
18	293.8	3.7	188207	2	AC021800	AC021800 Homo sapi
19	292	3.6	193126	9	AL139275	AL139275 Human DNA
20	290.4	3.6	158983	2	AC073070	AC073070 Homo sapi
21	289.6	3.6	150145	9	AC069216	AC069216 Homo sapi
22	289.6	3.6	178168	2	AC024412	AC024412 Homo sapi
23	289	3.6	126394	9	AC109998	AC109998 Homo sapi
24	289	3.6	194860	9	AC091903	AC091903 Homo sapi
25	288.2	3.6	65608	9	AL450243	AL450243 Human DNA
26	288.2	3.6	181618	2	AL442129	AL442129 Homo sapi
27	288.2	3.6	190092	2	AC120502	AC120502 Lemur cat
28	284.4	3.5	147055	9	HS372K1	HS372K1 Human DNA
29	282.2	3.5	178905	9	AC024153	AC024153 Homo sapi
30	280.2	3.5	98232	9	AC012271	AC012271 Homo sapi
31	280.2	3.5	152005	9	AC100834	AC100834 Homo sapi
32	280	3.5	27724	9	EX284678	EX284678 Human DNA
33	276.2	3.4	140411	9	AL590392	AL590392 Human DNA
34	275.8	3.4	179986	9	AP001982	AP001982 Homo sapi
35	275.8	3.4	160800	2	AP002011	AP002011 Homo sapi
36	275.8	3.4	162584	2	AP002763	AP002763 Homo sapi
37	275.8	3.4	183674	2	AC016996	AC016996 Homo sapi
38	275.4	3.4	142533	9	BS000038	BS000038 Pan trogl
39	274	3.4	228353	2	AL445283	AL445283 Homo sapi
40	273.8	3.4	172461	2	AC083786	AC083786 Homo sapi
41	273.8	3.4	173010	9	CNS01DSU	AL122013 Human chr
42	271.6	3.4	154336	9	AC117409	AC117409 Homo sapi
43	271.6	3.4	154341	9	HS1164K7	AL592209 Homo sapi
44	268.8	3.3	204662	9	CNS01RHO	AL162191 Human chr
45	267.2	3.3	127790	9	AC008885	AC008885 Homo sapi

#### ALIGNMENTS

RESULT 1  
LOCUS AX451337 20300 bp DNA linear PAT 03-JUL-2002  
DEFINITION Sequence 3 from Parent W0021653.  
ACCESSION AX451337  
VERSION AX451337.1 GI:21698388  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
artificial sequences.  
REFERENCE  
AUTHORS Levitt,P.R., Minicis,K., Kodavali,V.C. and Nimgaonkar,V.L.  
TITLE Methods and systems for facilitating the diagnosis and treatment of schizophrenia  
JOURNAL Patent: WO 021653-A 3 28-FEB-2002;

FEATURES  
Source University of Pittsburgh (US)  
Location/Qualifiers  
1..20300  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="A genomic sequence containing RGS4 nucleic acid  
sequence and sequences upstream and downstream to the  
RGS4 nucleic acid sequence"

ORIGIN

Query Match 100.0%; Score 8041; DB 6; Length 20300;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 8041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAATTTGAAGTCTGGGACGTTACAAAATAGAGTCAATAAGTTTACCTTAATTTACC 60  
Db 4110 CAATTTGAAGTCTGGGACGTTACAAAATAGAGTCAATAAGTTTACCTTAATTTACC 4169

QY 61 AAGATTTCTAGAGGATCTATATTTGGAATCCAGATTCGCTCTCTTAAGTTCAAGCA 120  
Db 4170 AAGATTTCTAGAGGATCTATATTTGGAATCCAGATTCGCTCTCTTAAGTTCAAGCA 4229

QY 121 CTTTCATGACACCACTACTGTTCTTCCACCTGACATGCAATGAACTTATGAAA 180  
Db 4230 CTTTCATGACACCACTACTGTTCTTCCACCTGACATGCAATGAACTTATGAAA 4289

QY 181 CTGCTGTTCTATCTCTGGGCTAAATGTTGCAAAAAAATTTAATCTTTGGATTAAGC 240  
Db 4290 CTGCTGTTCTATCTCTGGGCTAAATGTTGCAAAAAAATTTAATCTTTGGATTAAGC 4349

QY 241 TATTTGGGTTTCTCTACTCTTGGGAAACAAGTTTCTCCCTGGCTAATTAAGT 300  
Db 4350 TATTTGGGTTTCTCTACTCTTGGGAAACAAGTTTCTCCCTGGCTAATTAAGT 4409

QY 301 GTGGTATGTTCTTCCAGGAAATCAATGATGATCACTGCTGCTCAATGTCAGGG 360  
Db 4410 GTGGTATGTTCTTCCAGGAAATCAATGATGATCACTGCTGCTCAATGTCAGGG 4469

QY 361 TTGAGTCTCTGATTTATGACATGAGCCCAACAAGCTTGAGCAAAAGATTGACAT 420  
Db 4470 TTGAGTCTCTGATTTATGACATGAGCCCAACAAGCTTGAGCAAAAGATTGACAT 4529

QY 421 TCCCAAAAGTAAGACATCTGGAAGTCCCTGTTTACCTTCTGTAATACAGATCTCC 480  
Db 4530 TCCCAAAAGTAAGACATCTGGAAGTCCCTGTTTACCTTCTGTAATACAGATCTCC 4589

QY 481 AGCCCATATCTTTGTTTGTAGTCTTAAATCAATACTGAATCTCATATGATGCTA 540  
Db 4590 AGCCCATATCTTTGTTTGTAGTCTTAAATCAATACTGAATCTCATATGATGCTA 4649

QY 541 GGCATTTAGTAAACAATAAAGAGAGGAGGCTTCTGACAATGAGAGGAATGTC 600  
Db 4650 GGCATTTAGTAAACAATAAAGAGAGGAGGCTTCTGACAATGAGAGGAATGTC 4709

QY 601 ATCTGAAGTGTGCAAGCACAGCTCGGAGCTTGAGCTTGGCTTACATCTGCCAGTGG 660  
Db 4710 ATCTGAAGTGTGCAAGCACAGCTCGGAGCTTGAGCTTGGCTTACATCTGCCAGTGG 4769

QY 661 AGGATCAGTCCCCCATTTAATCATCTGTAGAACTAAAGACGAAAGCTGCCCAATGA 720  
Db 4770 AGGATCAGTCCCCCATTTAATCATCTGTAGAACTAAAGACGAAAGCTGCCCAATGA 4829

QY 721 CTATTTCCCTGCAATTTGATACCGTCAATCTTGAGAAATGTTTCTTTGTTCTCCCTG 780  
Db 4830 CTATTTCCCTGCAATTTGATACCGTCAATCTTGAGAAATGTTTCTTTGTTCTCCCTG 4889

QY 781 AGCAAGGTTGAAAAATTTGAATTTAATCTAGAGACACACATAGTTCAATCTGCTG 840  
Db 4890 AGCAAGGTTGAAAAATTTGAATTTAATCTAGAGACACACATAGTTCAATCTGCTG 4949

QY 841 TGTGCTGAATGTCTGCCCCAGTAGAGAAACAGTTTCTTAAGCCTATGTCAACAT 900  
Db 841 TGTGCTGAATGTCTGCCCCAGTAGAGAAACAGTTTCTTAAGCCTATGTCAACAT 900

Db 4950 TGTGCTGAATGTCTGCCCCAGTAGAGAAACAGTTTCTTAAGCCTATGTCAACAT 5009

QY 901 ACCCTTCAGATGTTAGATTTTACAAATTAAGAACTTAAATAGCTTCAAACTTTTG 960  
Db 5010 ACCCTTCAGATGTTAGATTTTACAAATTAAGAACTTAAATAGCTTCAAACTTTTG 5069

QY 961 CCAGTTTCTGATATCCAAATCTATCTTTACTCTGCTCCCAAGCTTTCTTAGAA 1020  
Db 5070 CCAGTTTCTGATATCCAAATCTATCTTTACTCTGCTCCCAAGCTTTCTTAGAA 5129

QY 1021 TGCTAACCTGATCGGCTTAAGTACTTAAGTAACTCTCTCCCAATTAATCTACAGTA 1080  
Db 5130 TGCTAACCTGATCGGCTTAAGTACTTAAGTAACTCTCTCCCAATTAATCTACAGTA 5189

QY 1081 AATTCGTCTTCAGAGTAACAGAAACACCTTTAGTTCTGACATATTCGTGACCTT 1140  
Db 5190 AATTCGTCTTCAGAGTAACAGAAACACCTTTAGTTCTGACATATTCGTGACCTT 5249

QY 1141 CATTTATCTCTCTCTCTCTCTCAAGCTGACATGAGGGGTGAACAGTGTATCATTTTC 1200  
Db 5250 CATTTATCTCTCTCTCTCTCTCAAGCTGACATGAGGGGTGAACAGTGTATCATTTTC 5309

QY 1201 TCTTCATCATTAAGGGTGCACCAACAAATCTCTATAGTAAAGACAGTTAATAGAGA 1260  
Db 5310 TCTTCATCATTAAGGGTGCACCAACAAATCTCTATAGTAAAGACAGTTAATAGAGA 5369

QY 1261 AAACCTAAACAAATTTATTTAATCAAGTTTACATGACATGAGAGTCTTCAGAAATGAG 1320  
Db 5370 AAACCTAAACAAATTTATTTAATCAAGTTTACATGACATGAGAGTCTTCAGAAATGAG 5429

QY 1321 ACCCAAGACCCAGGGGAACTGCTGTTTTTTTTGCTGAGTGTGATGAGATGATA 1380  
Db 5430 ACCCAAGACCCAGGGGAACTGCTGTTTTTTTTGCTGAGTGTGATGAGATGATA 5489

QY 1381 GCATGTAGCCATGTAGATTAGACAAAGATATGATCTAGTGTAAAGACTCAGGGGGA 1440  
Db 5490 GCATGTAGCCATGTAGATTAGACAAAGATATGATCTAGTGTAAAGACTCAGGGGGA 5549

QY 1441 AACACAGAGAGGCTGTCTATTCAGATCTCTGATCTCTCTCTATATGATGAT 1500  
Db 5550 AACACAGAGAGGCTGTCTATTCAGATCTCTGATCTCTCTCTCTATATGATGAT 5609

QY 1501 TCTTCTCTCTGATGTAGGAGACAGACTCTTCTTCATGAGAGGTCTTCAAGGAGAGAGG 1560  
Db 5610 TCTTCTCTCTGATGTAGGAGACAGACTCTTCTTCATGAGAGGTCTTCAAGGAGAGAGG 5669

QY 1561 AGAAGTGGCTTTTATGATTTTATAGCTTGTGCGGGAAGAGAGTCTATGTTCTAT 1620  
Db 5670 AGAAGTGGCTTTTATGATTTTATAGCTTGTGCGGGAAGAGAGTCTATGTTCTAT 5729

QY 1621 GACCCATCTTGGGGAGAGAAATTCGTTTCTGTGACTTGTCTTATGAGAAAGAGGA 1680  
Db 5730 GACCCATCTTGGGGAGAGAAATTCGTTTCTGTGACTTGTCTTATGAGAAAGAGGA 5789

QY 1681 GTAAGAGCAGAGAGGAGAGATGTCAGAAAGAGACTTGGCTTCTGAGGGCTTCC 1740  
Db 5790 GTAAGAGCAGAGAGGAGAGATGTCAGAAAGAGACTTGGCTTCTGAGGGCTTCC 5849

QY 1741 GCTCTGCTTATGATTCAGAGTACTTCTTATGATACCAAGAGACTATCTTTGGCATATG 1800  
Db 5850 GCTCTGCTTATGATTCAGAGTACTTCTTATGATACCAAGAGACTATCTTTGGCATATG 5909

QY 1801 TTTCTGAGCTCTAACTGCAATCATGCTAACTCTCTATGACTTCAAACTTCACT 1860  
Db 5910 TTTCTGAGCTCTAACTGCAATCATGCTAACTCTCTATGACTTCAAACTTCACT 5969

QY 1861 TGCTTTATCTTATGATGTTGATGATGATGATGATGATGATGATGATGATGATGATG 1920  
Db 5970 TGCTTTATCTTATGATGTTGATGATGATGATGATGATGATGATGATGATGATGATG 6029

QY 1921 ACTGCTGAGCTGGCATTAATTAATCAATTAATCAATTAATCAATTAATCAATTAAT 1980  
Db 6030 ACTGCTGAGCTGGCATTAATTAATCAATTAATCAATTAATCAATTAATCAATTAAT 6089

QY 1981 TTAAGTCTATGCTTCATCTGTAAATGAGATCATGAAAGACATTCCTCAGAGT 2040  
Db 6090 TTAAGTCTATGCTTCATCTGTAAATGAGATCATGAAAGACATTCCTCAGAGT 6149  
QY 2041 GATCATGAGAAATAGAGATTAAGTGGCATATAGTCTTAAACCGTGGCTTGTCTCAT 2100  
Db 6150 GATCATGAGAAATAGAGATTAAGTGGCATATAGTGGCTTAAACCGTGGCTTGTCTCAT 6209  
QY 2101 TAGTGAAGATTAATATCATCTGTATTAAGTGGCCATATGATGCTCTCTCTCT 2160  
Db 6210 TAGTGAAGATTAATATCATCTGTATTAAGTGGCCATATGATGCTCTCTCTCTCT 6269  
QY 2161 TTGTACAAGACTACATCTGTATTAATCATTTTGGGTCTCTCTGTGAAACCACTCCA 2220  
Db 6270 TTGTACAAGACTACATCTGTATTAATCATTTTGGGTCTCTCTGTGAAACCACTCCA 6329  
QY 2221 GATTCAAAAGATAGTATTAATCTATACCTGTGGTCTTCTGAGATTTGTAAAGTAA 2280  
Db 6330 GATTCAAAAGATAGTATTAATCTATACCTGTGGTCTTCTGAGATTTGTAAAGTAA 6389  
QY 2281 ATCTTCATTCATCATCATGATAGGTTCTTGAAACTCAACTTCAAGTGAAGACATTA 2340  
Db 6390 ATCTTCATTCATCATCATGATAGGTTCTTGAAACTCAACTTCAAGTGAAGACATTA 6449  
QY 2341 ACTAAACCAATTTTTTCTCATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2400  
Db 6450 ACTAAACCAATTTTTTCTCATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 6509  
QY 2401 TTCAAGGACCTGTGACCTGTCTTCACTTAAAGTCACTGTTCATATATCATTTGATG 2460  
Db 6510 TTCAAGGACCTGTGACCTGTCTTCACTTAAAGTCACTGTTCATATATCATTTGATG 6569  
QY 2461 ACATTTGAGGACTTACTAT 2520  
Db 6570 ACATTTGAGGACTTACTAT 6629  
QY 2521 CTGCTAACTGTGCTAACTGTCTCCCTGCTCCACACTGCGCCACTCATCTAGCTTT 2580  
Db 6630 CTGCTAACTGTGCTAACTGTCTCCCTGCTCCACACTGCGCCACTCATCTAGCTTT 6689  
QY 2581 CTTTCACAAGAGTCAAGATGATCAGATGAGACCCCTCTGCTCTGTTCTTCATGG 2640  
Db 6690 CTTTCACAAGAGTCAAGATGATCAGATGAGACCCCTCTGCTCTGTTCTTCATGG 6749  
QY 2641 ATTTCGACCTGCACTGTGTAAGTCAAGCCTCTGACACAGCCTTACAAATCTGTGACG 2700  
Db 6750 ATTTCGACCTGCACTGTGTAAGTCAAGCCTCTGACACAGCCTTACAAATCTGTGACG 6809  
QY 2701 ATCTATCGTTTACTTTCCATCTCTTTTATGCTACTTTCATCTGTTCTCAATCTCTTA 2760  
Db 6810 ATCTATCGTTTACTTTCCATCTCTTTTATGCTACTTTCATCTGTTCTCAATCTCTTA 6869  
QY 2761 GCTATGCTGGCCCTCTTCTTCTTCCCATTTTTTTTAAATTTGTATATA 2820  
Db 6870 GCTATGCTGGCCCTCTTCTTCTTCCCATTTTTTTTAAATTTGTATATA 6929  
QY 2821 TTATGAGGTTATAGTAAATCTTTTATGATGATAGTGTATAGTAAATATAGAG 2880  
Db 6930 TTATGAGGTTATAGTAAATCTTTTATGATGATAGTGTGTATAGTAAATATAGAG 6989  
QY 2881 CTTTATGAGGTTATCATCACTGATGATGATCATTTGATGATGATGATGATGATGATGAT 2940  
Db 6990 CTTTATGAGGTTATCATCACTGATGATGATGATCATTTGATGATGATGATGATGATGAT 7049  
QY 2941 TCGGCTGATCTTCTGCCCCCTGGGATTCATCACTGATGATGATGATGATGATGATGAT 3000  
Db 7050 TCGGCTGATCTTCTGCCCCCTGGGATTCATCACTGATGATGATGATGATGATGATGAT 7109  
QY 3001 AGTAATTTCTCAACATCGGCTGATCTTCTGCCCCCTGGGATTCATCACTGATGATGAT 3060  
Db 7110 AGTAATTTCTCAACATCGGCTGATCTTCTGCCCCCTGGGATTCATCACTGATGATGAT 7169

QY 3061 GCATTTGACCCCTTAAAGTAATTTCTCACATCGGCTGACTCTTGGCCCCCTATCTCTCT 3120  
Db 7170 GCATTTGACCCCTTAAAGTAATTTCTCACATCGGCTGACTCTTGGCCCCCTATCTCTCT 7229  
QY 3121 GAGGCTCCATTTTCCATTCATTCACACTGATCATATGATGATGATGATGATGATGATGAT 3180  
Db 7230 GAGGCTCCATTTTCCATTCATTCACACTGATCATATGATGATGATGATGATGATGATGAT 7289  
QY 3181 ACTTATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240  
Db 7290 ACTTATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7349  
QY 3241 AATGGCCCCAGTTTCATCTAGGCTGCTGCAAAAGGACATGATTTCACTTTTATAGG 3300  
Db 7350 AATGGCCCCAGTTTCATCTAGGCTGCTGCAAAAGGACATGATTTCACTTTTATAGG 7409  
QY 3301 TATGTTCTTCCCAATTTAGATTAAGTAAGACATGCACTTGTCTTACTTCTATTTGGAAT 3360  
Db 7410 TATGTTCTTCCCAATTTAGATTAAGTAAGACATGCACTTGTCTTACTTCTATTTGGAAT 7469  
QY 3361 ACTTATTTCTAGGCTTCTGATTTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3420  
Db 7470 ACTTATTTCTAGGCTTCTGATTTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7529  
QY 3421 TACCACTCTTCAAAAGGAGGCTTCTGACACACTTGGCTGATTTAGCCCTTACCATCT 3480  
Db 7530 TACCACTCTTCAAAAGGAGGCTTCTGACACACTTGGCTGATTTAGCCCTTACCATCT 7589  
QY 3481 GATTACTCTCTAGACATCACTCTGCCATTTTATATGATGATGATGATGATGATGATGAT 3540  
Db 7590 GATTACTCTCTAGACATCACTCTGCCATTTTATATGATGATGATGATGATGATGATGAT 7649  
QY 3541 CACCTGATTTGTTATTTTCTGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3600  
Db 7650 CACCTGATTTGTTATTTTCTGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7709  
QY 3601 ATTTATCTGCTGTATCAGGTACTGCTTCAAAACAGACCTGATTAAGAGTGTGCTGA 3660  
Db 7710 ATTTATCTGCTGTATCAGGTACTGCTTCAAAACAGACCTGATTAAGAGTGTGCTGA 7769  
QY 3661 AAAGTATTTCTTAAACAAATGAAACAAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 3720  
Db 7770 AAAGTATTTCTTAAACAAATGAAACAAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 7829  
QY 3721 CCACACTACAGAGTCACTTACAGATGACCAATCATAGATTAAGATTTGATGCTTTG 3780  
Db 7830 CCACACTACAGAGTCACTTACAGATGACCAATCATAGATTAAGATTTGATGCTTTG 7889  
QY 3781 TGCATATCAGAAAGAGTGAAGTGTCAAGTATATATGATGATGATGATGATGATGATGAT 3840  
Db 7890 TGCATATCAGAAAGAGTGAAGTGTCAAGTATATATGATGATGATGATGATGATGATGAT 7949  
QY 3841 TCTCCTAAACCAATCTGCAAGTCCCAAGTGAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 3900  
Db 7950 TCTCCTAAACCAATCTGCAAGTCCCAAGTGAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 8009  
QY 3901 TTACAAACATTTTAAAGTCTTAAACTCTTAAGAAAGCAAGATTTATGTCGAAG 3960  
Db 8010 TTACAAACATTTTAAAGTCTTAAACTCTTAAGAAAGCAAGATTTATGTCGAAG 8069  
QY 3961 AAGTGAAGTAAAGTAAAGTCTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4020  
Db 8070 AAGTGAAGTAAAGTAAAGTCTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 8129  
QY 4021 GTTAGTAAATATGCTTGAAGAGAGAGCTCTGCTTCAAGAGCAAGAACAAATTTGG 4080  
Db 8130 GTTAGTAAATATGCTTGAAGAGAGAGCTCTGCTTCAAGAGCAAGAACAAATTTGG 8189  
QY 4081 CAAGTGAACATCTCTGATTAATTAATGATTAATTTTGGGTTATGATGATGATGATGAT 4140  
Db 8190 CAAGTGAACATCTCTGATTAATTAATGATTAATTTTGGGTTATGATGATGATGATGAT 8249  
QY 4141 GTATATAGATGATTAATAAATTTATTTGATGATGATGATGATGATGATGATGATGAT 4200

Db 8250 GATTAATAGATATAAATAATATTTGACTAGACACTTGTAGATTAGAAATATCTAATT 8309  
QY 4201 TACACAGTTTACCTTAATTTGATAGAGACTGTGATGATGGGATAGACATGATGCAATCC 4260  
Db 8310 TACACAGTTTACCTTAATTTGATAGAGACTGTGATGATGGGATAGACATGATGCAATCC 8369  
QY 4261 ACAATACTGATGATGAGACACCTGTATCTGAGACCACTCTGTATGATAGAGACTGAA 4320  
Db 8370 ACAATACTGATGATGAGACACCTGTATCTGAGACCACTCTGTATGATAGAGACTGAA 8429  
QY 4321 CCAAGAGAGTCACTTCTCTTCTGGGCTCTATTTCTTTGGTGAATGAGAGGT 4380  
Db 8430 CCAAGAGAGTCACTTCTCTTCTGGGCTCTATTTCTTTGGTGAATGAGAGGT 8489  
QY 4381 TAGGTAAGATGCTTTGAGTCCCATTTTGTCTTTAAAGTCCATCTATTTGAGAT 4440  
Db 8490 TAGGTAAGATGCTTTGAGTCCCATTTTGTCTTTAAAGTCCATCTATTTGAGAT 8549  
QY 4441 TATATTTAACTGATGACAAATCAGGCTCTCTTATTTCTAAAGTCCATCTATTTGAGAT 4500  
Db 8550 TATATTTAACTGATGACAAATCAGGCTCTCTTATTTCTAAAGTCCATCTATTTGAGAT 8609  
QY 4501 ATTGGAATTTTCCAGGATCAGTAAATCTTTTGGTACTGACTATTTCTGAAATCA 4560  
Db 8610 ATTGGAATTTTCCAGGATCAGTAAATCTTTTGGTACTGACTATTTCTGAAATCA 8669  
QY 4561 ATCTATTTAGATGATGATCTCTTTTGGTGGCCAGGTAAACAAGAGGCCATGCTTTTC 4620  
Db 8670 ATCTATTTAGATGATGATCTCTTTTGGTGGCCAGGTAAACAAGAGGCCATGCTTTTC 8729  
QY 4621 TTGAGTGAACCTTCTTCCCTTTTATTTAGTCTGACCTCTTAAATGATGATGATGATGAT 4680  
Db 8730 TTGAGTGAACCTTCTTCCCTTTTATTTAGTCTGACCTCTTAAATGATGATGATGATGAT 8789  
QY 4681 TTCAATTTCCCTGGTCACTTCTCTTGGTCTGAGGGCTTCTAGTTTCAATTTGACCTTC 4740  
Db 8790 TTCAATTTCCCTGGTCACTTCTCTTGGTCTGAGGGCTTCTAGTTTCAATTTGACCTTC 8849  
QY 4741 AGTTCTTCCACACCAACATCAAGAGTGGCTGCAACATTTGATTTGATTTATTTAT 4800  
Db 8850 AGTTCTTCCACACCAACATCAAGAGTGGCTGCAACATTTGATTTGATTTATTTATTTAT 8809  
QY 4801 TCAAGGAAAAGTTGCCAGTAGCTAATCCAAATTAATGCTCTTATGGGCGCTAGAGAC 4860  
Db 8910 TCAAGGAAAAGTTGCCAGTAGCTAATCCAAATTAATGCTCTTATGGGCGCTAGAGAC 8969  
QY 4861 TTTTTCCTAATTTTAAATGATCTTCTGTAGATTAATGATGCTTCCCTCCACCACTTTACATTT 4920  
Db 8970 TTTTTCCTAATTTTAAATGATCTTCTGTAGATTAATGATGCTTCCCTCCACCACTTTACATTT 9029  
QY 4921 GTCTGCTGTCTCTTCTGCTGCTAGTCAATGAGACGTGTGTAGTGGGCGCAAGTGGGA 4980  
Db 9030 GTCTGCTGTCTCTTCTGCTGCTAGTCAATGAGACGTGTGTAGTGGGCGCAAGTGGGA 9089  
QY 4981 TGTTCAAAGGACGATTTGGGTAGGGGCAATATGGGCAATGCTTTGTGCCATTTCTTCT 5040  
Db 9090 TGTTCAAAGGACGATTTGGGTAGGGGCAATATGGGCAATGCTTTGTGCCATTTCTTCT 9149  
QY 5041 ATATTTTGGTATTTTGCATCTCACTGAGACCACTAATTTTTCATCTCTTCCACCTAA 5100  
Db 9150 ATATTTTGGTATTTTGCATCTCACTGAGACCACTAATTTTTCATCTCTTCCACCTAA 9209  
QY 5101 CTAATTTGATGCTCTGTCTTTTATATATTAAGATAGCTCACTGTAGCTATGATGAGGA 5160  
Db 9210 CTAATTTGATGCTCTGTCTTTTATATATTAAGATAGCTCACTGTAGCTATGATGAGGA 9269  
QY 5161 ACCCTATGCTTTCTTAATGAAAGCTTTTGGTCAAGTATGAGAAATTAATTTCCCTTT 5220  
Db 9270 ACCCTATGCTTTCTTAATGAAAGCTTTTGGTCAAGTATGAGAAATTAATTTCCCTTT 9329  
QY 5221 CCACTTATAGCTTTCTCTGTAACCTGTGTAGGATTTGGTTTATGGCTATAGATGT 5280

Db 9330 CCACTTATAGCTTTCTCTGTAACCTGTGTAGGATTTGGTTTATGGCTATAGATGT 9389  
QY 5281 GAACACCTGATGATTTCTGTCCATGACGAGGATTTCACTTCAATGATTTGATGTAAG 5340  
Db 9390 GAACACCTGATGATTTCTGTCCATGACGAGGATTTCACTTCAATGATTTGATGTAAG 9449  
QY 5341 ATACTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5400  
Db 9450 ATACTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9509  
QY 5401 GAAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5460  
Db 9510 GAAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9569  
QY 5461 CTACATCTCTCTTCTTATCTTTTGAATATCAGTTTGGAGACTTGAATTAATGATGAT 5520  
Db 9570 CTACATCTCTCTTCTTATCTTTTGAATATCAGTTTGGAGACTTGAATTAATGATGAT 9629  
QY 5521 CATTAACCCATTTGATTAATGATTTGCTGTGATATTTGCTCATCTTATCTTCCCTTTGG 5580  
Db 9630 CATTAACCCATTTGATTAATGATTTGCTGTGATATTTGCTCATCTTATCTTCCCTTTGG 9689  
QY 5581 TCTCTGAGCTCCGCTTTAGACTTTTACAGAGCGAGCATGTGAAGAGAACTGTGA 5640  
Db 9690 TCTCTGAGCTCCGCTTTAGACTTTTACAGAGCGAGCATGTGAAGAGAACTGTGA 9749  
QY 5641 GTGCTAGGCTGAATTTCTGTGTTTACCAAGATTTCTGAAAGATTTCTGAGCAAGTTG 5700  
Db 9750 GTGCTAGGCTGAATTTCTGTGTTTACCAAGATTTCTGAAAGATTTCTGAGCAAGTTG 9809  
QY 5701 ATTACAGATTAAGAAATCTATTTTCTGAGGAGTTTGTATGCTGCGGCGCTATA 5760  
Db 9810 ATTACAGATTAAGAAATCTATTTTCTGAGGAGTTTGTATGCTGCGGCGCTATA 5869  
QY 5761 ACTGCTGTCACTCAGCTATTTTCTGACCTTCTTGTGATTAACAACCAAGCACT 5820  
Db 9870 ACTGCTGTCACTCAGCTATTTTCTGACCTTCTTGTGATTAACAACCAAGCACT 9929  
QY 5821 TCCGGAATCACTGCGTGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 5880  
Db 9930 TCCGGAATCACTGCGTGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 9989  
QY 5881 TCAAGTCTTGACCATGATTAATATGATGCTTCTTAATCCAAAGAGAAAGCATTTGGA 5940  
Db 9990 TCAAGTCTTGACCATGATTAATATGATGCTTCTTAATCCAAAGAGAAAGCATTTGGA 10049  
QY 5941 GTGAGCTTCTTAAGTAAGCTCCAGAAATTTCTGCTGTGATTTTCTTCCAGAAAGCACTT 6000  
Db 10050 GTGAGCTTCTTAAGTAAGCTCCAGAAATTTCTGCTGTGATTTTCTTCCAGAAAGCACTT 10109  
QY 6001 CTTGATTAATTTTTTTACAGCATATGATTAATTAATTTTGGAGCAATTTGATAC 6060  
Db 10110 CTTGATTAATTTTTTTACAGCATATGATTAATTAATTTTGGAGCAATTTGATAC 10169  
QY 6061 TTTTTCCTCTTTTCTAATAATTTTAAACCTCTGACATTTGTGTGAGCATTTGATAC 6120  
Db 10170 TTTTTCCTCTTTTCTAATAATTTTAAACCTCTGACATTTGTGTGAGCATTTGATAC 10229  
QY 6121 TTTCCCAATGCCCACTTTTCAAGAGATTTTCTGCTGCTCACTTAACATTTGCTAT 6180  
Db 10230 TTTCCCAATGCCCACTTTTCAAGAGATTTTCTGCTGCTCACTTAACATTTGCTAT 10289  
QY 6181 GGTGAGCTTTTCTTCTCATCTTTTCAAGGCTGTGAGAGGCAAGGAGACAGAGA 6240  
Db 10290 GGTGAGCTTTTCTTCTCATCTTTTCAAGGCTGTGAGAGGCAAGGAGACAGAGA 10349  
QY 6241 GGTGAGCTGTGAGAGGCTGTGATTTGGCTGGAAGGCTGTGAGGCTGTGATTAAGAG 6300  
Db 10350 GGTGAGCTGTGAGAGGCTGTGATTTGGCTGGAAGGCTGTGAGGCTGTGATTAAGAG 10409  
QY 6301 ACCCTTACAGGCTTACAGAGAGAGCTCAGAGATTTCTGACAAATCTTTTACCGGAGA 6360  
Db 10410 ACCCTTACAGGCTTACAGAGAGAGCTCAGAGATTTCTGACAAATCTTTTACCGGAGA 10469

QY	6361	GAGGCAAAAGTAACGCTCAAAAGCCGAACGACAGCTCTCTCGCGGCAATTCCTTCCTCGGCT	6420
Db	10470	GAGGCAAAAGTAACGCTCAAAAGCCGAACGACAGCTCTCTCGCGGCAATTCCTTCCTCGGCT	10523
QY	6421	GCGAATTCGAAGCTGTTAAATAAGATGTGCAAAAGGCGCTTCGAGTGTGCGGCTTCCTTTCG	6480
Db	10530	GCGAATTCGAAGCTGTTAAATAAGATGTGCAAAAGGCGCTTCGAGTGTGCGGCTTCCTTTCG	10589
QY	6481	TTGAGGAGGTAAAGTGGCTTTCAGCCATTAAACATATTTAACTTTGGCTAGACTTTC	6540
Db	10590	TTGAGGAGGTAAAGTGGCTTTCAGCCATTAAACATATTTAACTTTGGCTAGACTTTC	10649
QY	6541	AGTTATTTACATGTTGTTACTTACTTAACCTAGTTCGTGCAATTAGAAACAGTGTGCTCAG	6600
Db	10650	AGTTATTTACATGTTGTTACTTACTTAACCTAGTTCGTGCAATTAGAAACAGTGTGCTCAG	10709
QY	6601	GAGAGCAAGACTTTCCTTAACCTTTCCTCCAGACACAGTAGATATTTGACCTTAAGACATGT	6660
Db	10710	GAGAGCAAGACTTTCCTTAACCTTTCCTCCAGACACAGTAGATATTTGACCTTAAGACATGT	10769
QY	6661	GCTCCCAAAATTTGACGCCCTTAATGTGTGTTTGTGTGACCTCAGCTTTGGAACCTGTC	6720
Db	10770	GCTCCCAAAATTTGACGCCCTTAATGTGTGTTTGTGTGACCTCAGCTTTGGAACCTGTC	10829
QY	6721	TATCTTTAAGCCAGGCTCTAAGAAAGCTAGTTTAAATTAAGAAAGCAGATGAGGTTGAG	6780
Db	10830	TATCTTTAAGCCAGGCTCTAAGAAAGCTAGTTTAAATTAAGAAAGCAGATGAGGTTGAG	10889
QY	6781	GCTATGTACAGTGATCTGTAAATCTTCATCTGTGATTACTAGCTATTTAGACATCCC	6840
Db	10890	GCTATGTACAGTGATCTGTAAATCTTCATCTGTGATTACTAGCTATTTAGACATCCC	10949
QY	6841	TGAGTAGCATAGAAAGCTGGGCTCTGGGCTTCGTAGATGTATGCTACAACTTGTTCAGGA	6900
Db	10950	TGAGTAGCATAGAAAGCTGGGCTCTGGGCTTCGTAGATGTATGCTACAACTTGTTCAGGA	11009
QY	6901	AAGGTACCCAGATGAGGTTTGCTCTCATCATCAGAAAGGCCTATGCTTTCCGTGTGG	6960
Db	11010	AAGGTACCCAGATGAGGTTTGCTCTCATCATCAGAAAGGCCTATGCTTTCCGTGTGG	11069
QY	6961	TGCTGCAGTAACCTTCACTCTCTATGTTCTTTAAGCAAAATGTTAACAATGAGATAGAT	7020
Db	11070	TGCTGCAGTAACCTTCACTCTCTATGTTCTTTAAGCAAAATGTTAACAATGAGATAGAT	11129
QY	7021	TTTAAAGCCAGATCTTCCTTATCTCTCTGCCCCATCTTAGTTCTTGAAGTGTCTATAT	7080
Db	11130	TTTAAAGCCAGATCTTCCTTATCTCTCTGCCCCATCTTAGTTCTTGAAGTGTCTATAT	11189
QY	7081	GAGTTGGTTGAGAAATATGATCATTAACAATCAGTTAATAGTTTGTGAGAAGATCTCA	7140
Db	11190	GAGTTGGTTGAGAAATATGATCATTAACAATCAGTTAATAGTTTGTGAGAAGATCTCA	11249
QY	7141	TCTTAAAGCACTGTTTGTGTTAATATATCTGCCCTTAGTTTTTTAAAGACCTTACAGACA	7200
Db	11250	TCTTAAAGCACTGTTTGTGTTAATATATCTGCCCTTAGTTTTTTAAAGACCTTACAGACA	11309
QY	7201	TACAGCTATATTTGTTTGTGTTTGTTCGAAAAAGGTATTAAGAAATGECATTTCAGAGA	7260
Db	11310	TACAGCTATATTTGTTTGTGTTTGTTCGAAAAAGGTATTAAGAAATGECATTTCAGAGA	11369
QY	7261	AAGATCATATATTTAGCCAGTTGAAAAATTTAAACACAAATGAGTCATTTACATTCCTTA	7320
Db	11370	AAGATCATATATTTAGCCAGTTGAAAAATTTAAACACAAATGAGTCATTTACATTCCTTA	11429
QY	7321	ATCTTCGAGTCAAAAGGTAAAAAGTCAACTAAAGGTATTAAGTACTGCTGTTCTTATGCGAC	7380
Db	11430	ATCTTCGAGTCAAAAGGTAAAAAGTCAACTAAAGGTATTAAGTACTGCTTCTTATGCGAC	11489
QY	7381	TGCAAAATGAAATTTACCAAAATTTTATTTTGGAAATATATCTCAGAAACATATATTTTTT	7440
Db	11490	TGCAAAATGAAATTTACCAAAATTTTATTTTGGAAATATATCTCAGAAACATATATTTTTT	11549

QY	744	ATGTCCTTAAATCATTTCTTCCAAATATCTGTCACTTCAGAGATGGAAGTACG	7500
Db	11550	ATGTCATATTAATCAATTTACTTTCCAAATATCTGTCTTCAGAGATGGAAGTACG	1160
QY	7501	ATGCGCTTCTTTAAATGAGCAGAGAGGGCTGTGACAGAGATATCTATGAATTAAGTCTCT	7560
Db	11610	ATGCGCTTCTTTAAATGAGCAGAGAGGGCTGTGACAGAGATATCTATGAATTAAGTCTCT	1166
QY	7561	CTGACCTTCAGCGTTAATTTCTGTGAATGAGGTGAGCAATTAATCTCAAGCTTCACCTAA	7620
Db	11670	CTGACCTTCAGCGTTAATTTCTGTGAATGAGGTGAGCAATTAATCTCAAGCTTCACCTAA	1172
QY	7621	CTTGCAATATGAATGAACCCGTACAAAATATCAAGAGTGTACAGAGAAAATTATGCTCTGG	7680
Db	11730	CTTGCAATATGAATGAACCCGTACAAAATATCAAGAGTGTACAGAGAAAATTATGCTCTGG	1178
QY	7681	TAAATTTTTTGGAAAACGATATAAAGTAATTAATAAGCTCTGTCTCTCAAGAGTTAAGC	7740
Db	11790	TAAATTTTTTGGAAAACGATATAAAGTAATTAATAAGCTCTGTCTCTCAAGAGTTAAGC	1184
QY	7741	AGCTAATCTTAAGAGGTAACTCTATGTACAGACAGATGAAGTCTCTTCCCTTCTCTCT	7800
Db	11850	AGCTAATCTTAAGAGGTAACTCTATGTACAGACAGATGAAGTCTCTTCCCTTCTCTCT	1190
QY	7801	CAATAAATTCGCAATCATCTATGTCCAAATCTTTTACCACAGTGCCTGAGGCTCCAGAGG	7860
Db	11910	CAATAAATTCGCAATCATCTATGTCCAAATCTTTTACCACAGTGCCTGAGGCTCCAGAGG	1196
QY	7861	AGCATATGCTCTTCAAGGTACATAGGTGTGGTGTGAGTTAGGACCAATCTAGAAATTC	7920
Db	11970	AGCATATGCTCTTCAAGGTACATAGGTGTGGTGTGAGTTAGGACCAATCTAGAAATTC	1202
QY	7921	CTGACTCCAGTAATCTTCGAAAGTCATTTTGTATTTTATATGTTTATATTAATGA	7980
Db	12030	CTGACTCCAGTAATCTTCGAAAGTCATTTTGTATTTTATATGTTTATATTAATGA	1208
QY	7981	TACTTCTTAAGACACTTACCCCTGTGATGATTAATACTCTAGAGATCTCAGGTGATC	8040
Db	12090	TACTTCTTAAGACACTTACCCCTGTGATGATTAATACTCTAGAGATCTCAGGTGATC	1214
QY	8041	C 8041	
Db	12150	C 12150	
RESULT 2			
AL583850			
LOCUS	AL583850	163229 bp	DNA linear PRI 15-NOV-2001
DEFINITION	Human DNA sequence from clone RP11-430G6 on chromosome 1, complete		
ACCESSION	AL583850		
VERSION	AL583850.5	GI:16973044	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Tracey, A.		
AUTHORS	Direct Submission		
TITLE	Submitted (15-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,		
JOURNAL	Cambridge, UK. E-mail enquiries: clonerequests@sanger.ac.uk		
COMMENT	On Nov 16, 2001 this sequence version replaced gi:15020514.		
	During differences assembly data is compared from overlapping clones.		
	Where differences are found these are annotated as variations		
	together with a note of the overlapping clone name. Note that the		
	variation annotation may not be found in the sequence submission		
	corresponding to the overlapping clone, as we submit sequences with		
	only a small overlap as described above.		
	This sequence was finished as follows unless otherwise noted: all		
	regions were either double-stranded or sequenced with an alternate		
	chemistry or covered by high quality data (1.e., phred quality >=		

30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Emi, EMBL; Swi, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone configs of human chromosome 1 constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chrl> Rpl1-43066 is from the library Rpl1-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: PBACe3.6

IMPORTANT: This sequence is not the entire insert of clone Rpl1-43066. It may be shorter because we sequence overlapping sections only once, except for a short overlap. The true right end of clone Rpl1-43066 is at 165329 in this sequence. The true right end of clone Rpl1-33182 is at 2000 in this sequence.

## FEATURES

## SOURCE

1..165329  
location=Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/clone="Rpl1-43066"  
/clone\_1b="RPl1-11.2"

## ORIGIN

## Query Match

Best Local Similarity 100.0%; Score 8039.4; DB 9; Length 165329;  
Matches 8040; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAAATTGAAGTCTCTGGGACGTTACAAAAGTAGAGTCATTAAGTTTACCTTATTTACC 60  
Db 75791 CAAATTGAAGTCTCTGGGACGTTACAAAAGTAGAGTCATTAAGTTTACCTTATTTACC 75850  
QY 61 AAGATTTCCTAGAGATCTATAATTTGAATCCAGATCTGCTCTCTGTAAGTTCAAGCA 120  
Db 75851 AAGATTTCCTAGAGATCTATAATTTGAATCCAGATCTGCTCTCTGTAAGTTCAAGCA 75910  
QY 121 CTTTCATGACACCACTGTTCTTCCACTGACCAATGCAATGAACTCTTATGAA 180  
Db 75911 CTTTCATGACACCACTGTTCTTCCACTGACCAATGCAATGAACTCTTATGAA 75970  
QY 181 CTGCTGTTTCTACTCTGGGCTAAATGTTGCAGAAAAAGATTATCTTTGGATAAGGC 240  
Db 75971 CTGCTGTTTCTACTCTGGGCTAAATGTTGCAGAAAAAGATTATCTTTGGATAAGGC 76030  
QY 241 TATTGGGTTTCTCTACTCTTGGGAAACAGGTTTCTTCCCTGGCTAATTAGT 300  
Db 76031 TATTGGGTTTCTCTACTCTTGGGAAACAGGTTTCTTCCCTGGCTAATTAGT 76090  
QY 301 GTGGATTGTTCTTCAGGGAATCAGTATGATCATCTGCTGCTATCAATGTCAAGG 360  
Db 76091 GTGGATTGTTCTTCAGGGAATCAGTATGATCATCTGCTGCTATCAATGTCAAGG 76150  
QY 361 TTGGAGTCTCTGATTATTGATGTCGCCACAAAGCTTGTGCAAGATTTGACACATT 420  
Db 76151 TTGGAGTCTCTGATTATTGATGTCGCCACAAAGCTTGTGCAAGATTTGACACATT 76210  
QY 421 TCCCAAAAGTAAGACATACTGGGAAGTCCCTGTTTACCTCTCTGTAATACAGATCTCC 480  
Db 76211 TCCCAAAAGTAAGACATACTGGGAAGTCCCTGTTTACCTCTCTGTAATACAGATCTCC 76270  
QY 481 AGCCCATATCTTTGCTTTTATAGTCCATAAAATCATATAGAACTCTCATTTGATGTA 540  
Db 76271 AGCCCATATCTTTGCTTTTATAGTCCATAAAATCATATAGAACTCTCATTTGATGTA 76330

QY 541 GGCATTGTAGTAACATTAAGAAGAGGAGGCTCTGCAACTGAGAGGAATTTGTC 600  
Db 76331 GGCATTGTAGTAACATTAAGAAGAGGAGGCTCTGCAACTGAGAGGAATTTGTC 76390  
QY 601 ATCTAAGTGTGCAAGACAGCTGGGGCTGAGCTTTGGCTTCAATCTTCCCAAGTGG 660  
Db 76391 ATCTAAGTGTGCAAGACAGCTGGGGCTGAGCTTTGGCTTCAATCTTCCCAAGTGG 76450  
QY 661 AGGATCAGGCCCATTTAATCATCTGGTAGACTAAAGAACGAAAGCCCTGCACATGA 720  
Db 76451 AGGATCAGGCCCATTTAATCATCTGGTAGACTAAAGAACGAAAGCCCTGCACATGA 76510  
QY 721 CTTATTCTCTGCAATTTGATAAGCTCAATCTTGAAGAAATGTTTCTTTCTCCCTG 780  
Db 76511 CTTATTCTCTGCAATTTGATAAGCTCAATCTTGAAGAAATGTTTCTTTCTCCCTG 76570  
QY 781 AGCAAGGTGGAAAAATTTGAATTTAAGTACAGACACATAGTTCATCTGCTG 840  
Db 76571 AGCAAGGTGGAAAAATTTGAATTTAAGTACAGACACATAGTTCATCTGCTG 76630  
QY 841 TGTGCTGAATGTCTGCCCCCAGTAGAAGAACAGTTCTTCTAAGCTTATTCACAAAT 900  
Db 76631 TGTGCTGAATGTCTGCCCCCAGTAGAAGAACAGTTCTTCTAAGCTTATTCACAAAT 76690  
QY 901 ACCTTCAGATGTAGCATTTTACATTTAAGGAACCTTAATAGCTTCAACCTTTTG 960  
Db 76691 ACCTTCAGATGTAGCATTTTACATTTAAGGAACCTTAATAGCTTCAACCTTTTG 76750  
QY 961 CCAGTCTCTGATATCCAAATCTATCTTTACTCTGCTCCCAAGCTTCTTCTAGAA 1020  
Db 76751 CCAGTCTCTGATATCCAAATCTATCTTTACTCTGCTCCCAAGCTTCTTCTAGAA 76810  
QY 1021 TGTACTGATCGGCTTAACTAATCTTCTCTCTCTCATTAATCTAGAGATA 1080  
Db 76811 TGTACTGATCGGCTTAACTAATCTTCTCTCTCTCATTAATCTAGAGATA 76870  
QY 1081 AATTCGCTCTCAGAGTAACAAAGAACACCCCTTAGTTCTCAGCATTTCCGCACTT 1140  
Db 76871 AATTCGCTCTCAGAGTAACAAAGAACACCCCTTAGTTCTCAGCATTTCCGCACTT 76930  
QY 1141 CATTTATCTCTCTCTCTCTCAAGCTGAGTAGGGGTGAAAAAGTGTATACATTTTC 1200  
Db 76931 CATTTATCTCTCTCTCTCTCAAGCTGAGTAGGGGTGAAAAAGTGTATACATTTTC 76990  
QY 1201 TCTTCATCATAAAGGTGCAACCAAACTCTATAGTAAAGACAGTTAATAAGAGA 1260  
Db 76991 TCTTCATCATAAAGGTGCAACCAAACTCTATAGTAAAGACAGTTAATAAGAGA 77050  
QY 1261 AAACCTAACAAATTTATTATCAATCAATGATGATGAGAGTCTTCAGAAATGAAG 1320  
Db 77051 AAACCTAACAAATTTATTATCAATCAATGATGATGAGAGTCTTCAGAAATGAAG 77110  
QY 1321 ACCCAAGACCCAGGGGAAATCTGTTTTTTTGTAGGTTGCATGAAGATGAGATA 1380  
Db 77111 ACCCAAGACCCAGGGGAAATCTGTTTTTTTGTAGGTTGCATGAAGATGAGATA 77170  
QY 1381 GCATGTAGCCATGTAGATTAGACAAAAGATATGATCTAGTGTAAAGACTCAGGGGA 1440  
Db 77171 GCATGTAGCCATGTAGATTAGACAAAAGATATGATCTAGTGTAAAGACTCAGGGGA 77230  
QY 1441 AACACAGAGAGGCTGTCTATATCAATCTTCTTATGATCTCTCTCTATATATGAT 1500  
Db 77231 AACACAGAGAGGCTGTCTATATCAATCTTCTTATGATCTCTCTCTATATATGAT 77290  
QY 1501 TCTTCTCTCTGAGTATGAGGAGAGACTCTTCTTCAATGAGGTTCTCAAGGAGAGAGG 1560  
Db 77291 TCTTCTCTCTGAGTATGAGGAGAGACTCTTCTTCAATGAGGTTCTCAAGGAGAGAGG 77350  
QY 1561 AGAAAGTGGCTTTTATGATTTATGAGCTTGTGAGGAGAGAGGATTTATGATTTAT 1620  
Db 77351 AGAAAGTGGCTTTTATGATTTATGAGCTTGTGAGGAGAGAGGATTTATGATTTAT 77410  
QY 1621 GACCATCTTTGGGGAAGAGAAATCTGTTCTGTGACTTGTCTATGAAGAAAGAGA 1680



[illegible]

Db	78491	ATCTATGCTTACTTTTTCGACATCGCTTTTATGCTATCTTCACTTGTCTCAATCTCTTA	78555
Qy	2761	GCTATGCTGAGCCCTCTGTGCTCTTCCATTTTTTTTAAATTTTAAATTTGTATATA	2820
Db	78551	GCTATGCTGAGCCCTCTGTGCTCTTCCATTTTTTTTAAATTTTAAATTTGTATATA	78610
Qy	2821	TTTATGGGTATATAGTAAGTAATCTTTTATAGTCATAGGTGATATAGTATATAATACAGG	2880
Db	78611	TTTATGGGTATATAGTAAGTAATCTTTTATAGTCATAGGTGATATAGTATATAATACAGG	78670
Qy	2881	CTTTTAGGCTATTCATCACTGATGATGTATCATTTGTACCCCTTAATTTCTACCA	2940
Db	78671	CTTTTAGGCTATTCATCACTGATGATGTATCATTTGTACCCCTTAATTTCTACCA	78730
Qy	2941	TCCGCTGACTCTTGGCCCCCTGGGATTTGATACCGGAATGATGTAAGTAAGCCCTTA	3000
Db	78731	TCCGCTGACTCTTGGCCCCCTGGGATTTGATACCGGAATGATGTAAGTAAGCCCTTA	78790
Qy	3001	AGTAATTTTCAACCATCCGTGACTTTTGGCCCCCTGGGATTTGATACCGGAATGATGT	3060
Db	78791	AGTAATTTTCAACCATCCGTGACTTTTGGCCCCCTGGGATTTGATACCGGAATGATGT	78850
Qy	3061	GCATTTGACCCCTTAAGTATTTTCTCACATCCGCTGATCTTTGCCCCCTCATCTCT	3120
Db	78851	GCATTTGACCCCTTAAGTATTTTCTCACATCCGCTGATCTTTTCCCCCTCATCTCTCT	78910
Qy	3121	GAGGCTGCATTTGCATATTCATTCACACCTCTACATCATATGTATGACATATTTAGCTCT	3180
Db	78911	GAGGCTGCATTTGCATATTCATTCACACCTCTACATCATATGTATGACATATTTAGCTCT	78970
Qy	3181	ACTTATATAGTATACATGCAATATTTGTCTTTCTGTCTGTCTTTGTTTACTATATGAT	3240
Db	78971	ACTTATATAGTATACATGCAATATTTGTCTTTCTGTCTGTCTTTGTTTACTATATGAT	79030
Qy	3241	AATGGCCCCCAGTTCTATATCTAGGCTCTGCAAAAGCATGATTTGATCTTTTATGAGC	3300
Db	79031	AATGGCCCCCAGTTCTATATCTAGGCTCTGCAAAAGCATGATTTGATCTTTTATGAGC	79090
Qy	3301	TATGTTCTTTCCCAATTTAGATATAGAACAACATCGCATTTGCTCTTACTTATTTTGAAT	3360
Db	79091	TATGTTCTTTCCCAATTTAGATATAGAACAACAACATCGCATTTGCTCTTACTTATTTGAAT	79150
Qy	3361	ACTATATCTTCAGGCTTTGCAATGTTCTTCTCTCTCAACCCATGAATAATCTCATTTTGA	3420
Db	79151	ACTATATCTTCAGGCTTTGCAATGTTCTTCTCTCTCAACCCATGAATAATCTCATTTTGA	79210
Qy	3421	TACACCTCTTTCAAAGAGGCTTTCTGTGACCACTTGGCTGATTAAGCCCTTCAACATCT	3480
Db	79211	TACACCTCTTTCAAAGAGGCTTTCTGTGACCACTTGGCTGATTAAGCCCTTCAACATCT	79270
Qy	3481	GATTACTCTCTAGACATCACTCTGCCATTTTATATGTATAGGTCAAAATCTTGAAAT	3540
Db	79271	GATTACTCTCTAGACATCACTCTGCCATTTTATATGTATAGGTCAAAATCTTGAAAT	79330
Qy	3541	CACCTGATTTTGTATTTATTTCTGACTCCTCTACTAGATGAAGAACTCTACTAGAGGAG	3600
Db	79331	CACCTGATTTTGTATTTATTTCTGACTCCTCTACTAGATGAAGAACTCTACTAGAGGAG	79390
Qy	3601	ATTTTATCTCTGTATCAAGTACTGCTTCAACACACCTGTATACAGATAGGTGTCTA	3660
Db	79391	ATTTTATCTCTGTATCAAGTACTGCTTCAACACACCTGTATACAGATAGGTGTCTA	79450
Qy	3661	AAATATTTCTTAAACAAATGTAACAAATPAAAGAGTATCTTTGAGAGTAACTCTT	3720
Db	79451	AAATATTTCTTAAACAAATGTAACAAATPAAAGAGTATCTTTGAGAGTAACTCTT	79510
Qy	3721	CCACATTCACAGAGTCTTCAAGAAATGACAAATCATAGAAATTAACAGAAATTTGATCTTTG	3780
Db	79511	CCACATTCACAGAGTCTTCAAGAAATGACAAATCATAGAAATTAACAGAAATTTGATCTTTG	79570
Qy	3781	TGCATATTCAGAGAAAGAGTGTGAAGTTGTCAAGTATCATAGATGTAACATCTCTGGC	3840
Db	79571	TGCATATTCAGAGAAAGAGTGTGAAGTTGTCAAGTATCATAGATGTAACATCTCTGGC	79630

QY	3841	TTCTCAACCAATCTTCGACAGTCCCAAGTGAAAAAGTAAGTTAACTCATGTGAAGCTT	3900
Db	79631	TTCTCAAAACCAATCTTCGACAGTCCCAAGTGAAAAAGTAAGTTAACTCATGTGAAGCTT	79630
QY	3901	TTACAAACACTTTTAAAAAGCTCTAAATCCTTAAGAAAGCAAGATTTAATAGTCAAG	3960
Db	79691	TTACAAACACTTTTAAAAAGCTCTAAATCCTTAAGAAAGCAAGATTTAATAGTCAAG	79750
QY	3961	AAGGAGTAAACATGAAGAATGCTTCGACAGAGTAAATGAGCTAACCAAAAGTTAGACAT	4020
Db	79751	AAGGAGTAAACATGAAGAATGCTTCGACAGAGTAAATGAGCTAACCAAAAGTTAGACAT	79810
QY	4021	GTTAGTAAATATGTCTTGAAGAGCAGCTCCGCTTTCAAGGAGCAAAACAAATTGGG	4080
Db	79811	GTTAGTAAATATGTCTTGAAGAGCAGCTCCGCTTTCAAGGAGCAAAACAAATTGGG	79870
QY	4081	CAAGTGAACACTCCTTGATTAATAANTGTGTAAATTAAATTTGGGTTATGTTCTATCTGT	4140
Db	79871	CAAGTGAACACTCCTTGATTAATAANTGTGTAAATTAAATTTGGGTTATGTTCTATCTGT	79930
QY	4141	GTAATAAAGATGATATAAATTTATTTACTAGACACTTGTAGTTTGAATAATCTTAT	4200
Db	79931	GTAATAAAGATGATATAAATTTATTTACTAGACACTTGTAGTTTGAATAATCTTAT	79990
QY	4201	TACACAGTTTACCTTATTTGATATAGACTGTTGAGTATGGATAGCATGTGACAAATCC	4260
Db	79991	TACACAGTTTACCTTATTTGATATAGACTGTTGAGTATGGATAGCATGTGACAAATCC	80050
QY	4261	ACATPACTGAGTATCGAGACACCTGATCTGACCCAGCTCGTTGTAGAAAGCTGTAA	4320
Db	80051	ACATPACTGAGTATCGAGACACCTGATCTGACCCAGCTCGTTGTAGAAAGCTGTAA	80110
QY	4321	CCTCAGCAAGTCACTTTCCTTTCTGGGTCTATTTCCTTTGGTGAATGAGAGTGT	4380
Db	80111	CCTCAGCAAGTCACTTTCCTTTCTGGGTCTATTTCCTTTGGTGAATGAGAGTGT	80170
QY	4381	TAGGCTTGATTTGCCTTTGAGAGTCCCAATTTGCTTTAAAGTCCCAATCTATTCAGATGAT	4440
Db	80171	TAGGCTTGATTTGCCTTTGAGAGTCCCAATTTGCTTTAAAGTCCCAATCTATTCAGATGAT	80230
QY	4441	TATATTTAAGTCAATGACAAATCAGGCTTCCTTATTTCTAAGTGCAGACATATAAATT	4500
Db	80231	TATATTTAAGTCAATGACAAATCAGGCTTCCTTATTTCTAAGTGCAGACATATAAATT	80290
QY	4501	ATTGTGGAATTTCAAGGATCAGTAATCTTTTGGGTACTCATATGTTCTGTAAATCA	4560
Db	80291	ATTGTGGAATTTCAAGGATCAGTAATCTTTTGGGTACTCATATGTTCTGTAAATCA	80350
QY	4561	ATCATTTGAGTATCACTCTTTTAGGTGCGCAGTAAACAAAGAGGCATATGATCTTTC	4620
Db	80351	ATCATTTGAGTATCACTCTTTTAGGTGCGCAGTAAACAAAGAGGCATATGATCTTTC	80410
QY	4621	TTTGAGTGACCTTCTTCCCTTTTAATTAAGTCTGACCTCTTTAATGTACGTTCTGACTGA	4680
Db	80411	TTTGAGTGACCTTCTTCCCTTTTAATTAAGTCTGACCTCTTTAATGTACGTTCTGACTGA	80470
QY	4681	TTATATTTCCCTGTGCATCTTCCCTTGTGTGAGAGGCTTCAGTGTCAATATTCAGCTTC	4740
Db	80471	TTATATTTCCCTGTGCATCTTCCCTTGTGTGAGAGGCTTCAGTGTTCATATTCAGCTTC	80530
QY	4741	AGTTCTCTTCAACACCAATCAAGATGCGTGTCAACATTCATTTGTTCTATGTATTAAT	4800
Db	80531	AGTTCTCTTCAACACCAATCAAGATGCGTGTCAACATTCATTTGTTCTATGTATTAAT	80590
QY	4801	TCAAGGAAAAGTTGCCAGATGATATCCAAATAAGCCCTGTATGAGCGGCTAGAGAC	4860
Db	80591	TCAAGGAAAAGTTGCCAGATGATATCCAAATAAGCCCTGTATGAGCGGCTAGAGAC	80650
QY	4861	TTTTTCTATATTTAAATGCAATCTTCTGTGAATTAAGTCCCTCCACCACTTACATTT	4920
Db	80651	TTTTTCTATATTTAAATGCAATCTTCTGTGAATTAAGTCCCTCCACCACTTACATTT	80710

OY	4921	GTCGCTGTCTCCTTGCTCTCTGTCAATGAACGGTGTGTAGTAGCGGGCCAGTGTGGGA	4980
Dd	80711	GTCGTCTGTCTCCTTGCTCTCTGTCAATGAACGGTGTGTAGTAGCGGGCCAGTGTGGGA	80777
OY	4981	TGTTCAAGGGCACGTAATTGGGTAGGGCCACAATAAGGGCAATTGCCTTGTGCCAATCTTCTCT	5040
Dd	80771	TGTTCAAGGGCACGTAATTGGGTAGGGCCACAATAAGGGCAATTGCCTTGTGCCAATCTTCTCT	80831
OY	5041	ATATTTTTTGGTAATTTTGCATCTCACGTGGAAACCAACTAATTTTTCATCTCTTCACCTTAA	5100
Dd	80831	ATATTTTTTGGTAATTTTGCATCTCACGTGGAAACCAACTAATTTTTCATCTCTTCACCTTAA	80891
OY	5101	CATTTTGATGCTCTGTCTTCTTAATAATAAGTAATAGCTCACGTAGACCTAATGATCAGGA	5160
Dd	80891	CATTTTGATGCTCTGTCTTCTTAATAATAAGTAATAGCTCACGTAGACCTAATGATCAGGA	8095
OY	5161	ACCTATCTGCTTTCTAAATGAAGAAGCTGTTTGGTCAGATCTGACATTAATTTCCCTCTT	5220
Dd	80951	ACCTATCTGCTTTCTAAATGAAGAAGCTGTTTGGTCAGATCTGACATTAATTTCCCTCTT	81011
OY	5221	CAACTATATAGCTTCCCTCTGTACTCTGSGTGTAGAAATTTGGTTATAGGCTATAAGATGT	5280
Dd	81011	CAACTATATAGCTTCCCTCTGTACTCTGSGTGTAGAAATTTGGTTATAGGCTATAAGATGT	8107
OY	5281	GAACAACCTGGAATGTAATCTGTCCACGACGAGCATTTGAGTCAATGATTTGTATGTAAGA	5340
Dd	81071	GAACAACCTGGAATGTAATCTGTCCACGACGAGCATTTGAGTCAATGATTTGTATGTAAGA	8113
OY	5341	AATACGATTTGTGTAGAGTGTGTAGAAAACCTATATAGGGCTTAATTTCTTAACATCAGTT	5400
Dd	81131	AATACGATTTGTGTAGAGTGTGTAGAAAACCTATATAGGGCTTAATTTCTTAACATCAGTT	8119
OY	5401	GAAGGTGTGTATAGCAAGCAAACACTACATATTTTTCGCTGCTCTCTCTTTCTCT	5460
Dd	81191	GAAGGTGTGTATAGCAAGCAAACACTACATATTTTTCGCTGCTCTCTCTTTCTCT	8125
OY	5461	CTACATCTCTCTTTCTTTATCTTTTGAATATCAGTTTGAGACTTAATTAACATAGA	5520
Dd	81251	CTACATCTCTCTTTCTTTATCTTTTGAATATCAGTTTGAGACTTAATTAACATAGA	8131
OY	5521	CATTAACCCTATTGATATAAGAAATGCTGTGTATTTTCTCATCTACCTCCCTCTTGG	5580
Dd	81311	CATTAACCCTATTGATATAAGAAATGCTGTGTATTTTCTCATCTACCTCCCTCTTGG	8137
OY	5581	TCCTGAGCTGCCGGTTTGAACCTTTTACAGACGACGAGCATGTAGAGAGAAACTGTCA	5640
Dd	81371	TCCTGAGCTGCCGGTTTGAACCTTTTACAGACGACGAGCATGTAGAGAGAAACTGTCA	8143
OY	5641	GTCCTTAGGCTGAATTCTGTGTATACCAAGATTTCTTGAAABATATTTCTCGAGTCAGGTG	5700
Dd	81431	GTCCTTAGGCTGAATTCTGTGTATACCAAGATTTCTTGAAABATATTTCTCGAGTCAGGTG	8149
OY	5701	ATTACAGATATAGCAAAATCTAATTTTCTTAGGGGTAGTTTCTGTATGCTCCGGGCTTATA	5760
Dd	81491	ATTACAGATATAGCAAAATCTAATTTTCTTAGGGGTAGTTTCTGTATGCTCCGGGCTTATA	8155
OY	5761	ACTGTCTGTCAATCCAGCTATTTCTCTCCACCTTTCTTGTGTATACCAACCAAGCAACT	5820
Dd	81551	ACTGTCTGTCAATCCAGCTATTTCTCTCCACCTTTCTTGTGTATACCAACCAAGCAACT	8161
OY	5821	TCGCAAAATCACTGGTGTAGAACGANTATCTGCCAGCTCCCTTTTGGAAATCTGTAGGA	5880
Dd	81611	TCGCAAAATCACTGGTGTAGAACGANTATCTGCCAGCTCCCTTTTGGAAATCTGTAGGA	8167
OY	5881	TGAGATCTTGGACATGATATATATATGATGCTTAAATCCAAAGAGGAAAGCAATTGGGA	5940
Dd	81671	TGAGATCTTGGACATGATATATATATGATGCTTAAATCCAAAGAGGAAAGCAATTGGGA	8173
OY	5941	GTCACCTCTTAATATAGTCCAGAAATTCGCGCTGAGTCTTTCCCTCCAGGAACCACTT	6000
Dd	81731	GTCACCTCTTAATATAGTCCAGAAATTCGCGCTGAGTCTTTCCCTCCAGGAACCACTT	8179
OY	6001	CGTTATATATTTTTTTTTTACAGGCATATGAAATAAAACTAATTTTTTGCAGCAATTGTACAC	6060

Db 81791 CTTTATATATTTTTTTTTTACAGGCAATGAAATTAATAATATTTTGCACATTTGACAC 81850  
Qy 6061 TTTTTCCTTTCTTCTAGAAATTTCTAACCTCTGACATTTGGTGAAGACATTTGACATTTT 6120  
Db 81851 TTTTTCCTTTCTTCTAGAAATTTCTAACCTCTGACATTTGGTGAAGACATTTGACATTTT 81910  
Qy 6121 TTTTCCATATCCCTTCTTCTTCTGAGAAATTTTCTCTGCTGCTTCACTTAACTTTGCTGAT 6180  
Db 81911 TTTTCCATATCCCTTCTTCTTCTGAGAAATTTTCTCTGCTGCTTCACTTAACTTTGCTGAT 81970  
Qy 6181 GCGTCAGTCTTTTCTTCTCTCTCTCTCTTCTGAGGCTGAGAGGCGAGAGGAGAGAGAGGAG 6240  
Db 81971 GCGTCAGTCTTTTCTTCT 82030  
Qy 6241 GCTGCTATCTGACAGAGCGCT 6300  
Db 82031 GCTGCTATCTGACAGAGCGCT 82090  
Qy 6301 ACCCTTACAGGCTTCTGACAGAGAGAGCTCAGAGATTTCTGCAATATCTTTACCGAGAA 6360  
Db 82091 ACCCTTACAGGCTTCTGACAGAGAGAGCTCAGAGATTTCTGCAATATCTTTACCGAGAA 82150  
Qy 6361 GAGGCAAGATGAGCTCAAGCCGAAAGCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6420  
Db 82151 GAGGCAAGATGAGCTCAAGCCGAAAGCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 82210  
Qy 6421 GCGAATTCAGAGCTTAAATTAAGATGAGCAAGAGGCTGAGGCTCTCTCTCTCTCTCTCT 6480  
Db 82211 GCGAATTCAGAGCTTAAATTAAGATGAGCAAGAGGCTGAGGCTCTCTCTCTCTCTCTCT 82270  
Qy 6481 TTGAGAGGTAAGATTGCTTCTGAGCAATTAACAATTTAACTTTTGGCTAGACCTTTCTC 6540  
Db 82271 TTGAGAGGTAAGATTGCTTCTGAGCAATTAACAATTTAACTTTTGGCTAGACCTTTCTC 82330  
Qy 6541 AGTTATTTAAGATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6600  
Db 82331 AGTTATTTAAGATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 82390  
Qy 6601 GAGACGAGACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6660  
Db 82391 GAGACGAGACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 82450  
Qy 6661 GCTCCCAATTTCTGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6720  
Db 82451 GCTCCCAATTTCTGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 82510  
Qy 6721 TATTCCTTAAAGCAGTCTTAAAGAAAGTAGTTTAAATTTAAAGACGAGATGAGGTTTGA 6780  
Db 82511 TATTCCTTAAAGCAGTCTTAAAGAAAGTAGTTTAAATTTAAAGACGAGATGAGGTTTGA 82570  
Qy 6781 GCTATGATCAGTATCTGTAATATCTCATCTGATTTACTCTGTAATTTGAGCATCCC 6840  
Db 82571 GCTATGATCAGTATCTGTAATATCTCATCTGATTTACTCTGTAATTTGAGCATCCC 82630  
Qy 6841 TGGAGTACATAGAGCGCTGAGGCTTCTGATTTGATGCTTAACTTGTTCAGGA 6900  
Db 82631 TGGAGTACATAGAGCGCTGAGGCTTCTGATTTGATGCTTAACTTGTTCAGGA 82690  
Qy 6901 AAGGTACCCAGAAATGAGGTTGGCTTCATCATAGAAAGGCACTATGCTTCCGTG 6960  
Db 82691 AAGGTACCCAGAAATGAGGTTGGCTTCATCATAGAAAGGCACTATGCTTCCGTG 82750  
Qy 6961 TGGTGAAGTAATTTCACTCTCTATGCTTAAAGCAAAATGTTACATAGATATGAGT 7020  
Db 82751 TGGTGAAGTAATTTCACTCTCTATGCTTAAAGCAAAATGTTACATAGATATGAGT 82810  
Qy 7021 TTTAAAGCAGATCTTCTTCTTCTCTCTGCCCCATCTCTAGTTCTTAAGTGTCTCATAT 7080  
Db 82811 TTTAAAGCAGATCTTCTTCTTCTCTCTGCCCCATCTCTAGTTCTTAAGTGTCTCATAT 82870  
Qy 7081 GAGTTTGGTGAAGTAATTTGATCTTTCAAAATCAGTTAAAGTTTGTGAGAGATCTCA 7140

Db 82871 GAGTTTGGTGAAGTAATTTGATCTTACAAATCAGTTAATAGTTTGTGAGAGATCTCA 82930  
Qy 7141 TCTTAAAGCAGATCTTCTTCTTCTTCTGCTTGAATTTTCTTAAAGACCTTACAGACA 7200  
Db 82931 TCTTAAAGCAGATCTTCTTCTTCTTCTGCTTGAATTTTCTTAAAGACCTTACAGACA 82990  
Qy 7201 TACAGCATATCATTTGTTTGTGTTTGTTCAAAAAGGTAAAGAAATGACATTTAGAGA 7260  
Db 82991 TACAGCATATCATTTGTTTGTGTTTGTTCAAAAAGGTAAAGAAATGACATTTAGAGA 83050  
Qy 7261 AAGATCATATTTAGCCAGTTGAAATTTAAACAAATAGTGCATATTACTTA 7320  
Db 83051 AAGATCATATTTAGCCAGTTGAAATTTAAACAAATAGTGCATATTACTTA 83110  
Qy 7321 ATCTTGACGTCAAGGTAAAGTCAACCTTAAGGTATCTCTGCTTCTTAAAGCAC 7380  
Db 83111 ATCTTGACGTCAAGGTAAAGTCAACCTTAAGGTATCTCTGCTTCTTAAAGCAC 83170  
Qy 7381 TGCAAATAGAAATTTACCAAAATTTTATTTTGGAAATTAATCTGAGAAAAATTAATTTT 7440  
Db 83171 TGCAAATAGAAATTTACCAAAATTTTATTTTGGAAATTAATCTGAGAAAAATTAATTTT 83230  
Qy 7441 ATGTAATATTTAAACATTTACTTTTCCAAATTTCTGATTCAGAGATATGAGATATG 7500  
Db 83231 ATGTAATATTTAAACATTTACTTTTCCAAATTTCTGATTCAGAGATATGAGATATG 83290  
Qy 7501 ATGCTTCTTAAATAGACAGAGAGGCTGAGAGATATCTAATAATAGTTCT 7560  
Db 83291 ATGCTTCTTAAATAGACAGAGAGGCTGAGAGATATCTAATAATAGTTCT 83350  
Qy 7561 CTGACCTTCAAGCTTAAATTTTCTGAAATGAGGTGAGCAAAATTTCTTCAAGCTTCACTTAA 7620  
Db 83351 CTGACCTTCAAGCTTAAATTTTCTGAAATGAGGTGAGCAAAATTTCTTCAAGCTTCACTTAA 83410  
Qy 7621 CTGCAATATGAAATGAAACCGTACAAATAACAAAGGTGAGAGAAATTTATGCTCTG 7680  
Db 83411 CTGCAATATGAAATGAAACCGTACAAATAACAAAGGTGAGAGAAATTTATGCTCTG 83470  
Qy 7681 TAAATATTTTGCAGAAACAGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 7740  
Db 83471 TAAATATTTTGCAGAAACAGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 83530  
Qy 7741 AGCTAATCTAAGAGGTAATCTATATGCTGAGAGATGAACTGCTCTCTCTCTCTCTCTCTCT 7800  
Db 83531 AGCTAATCTAAGAGGTAATCTATATGCTGAGAGATGAACTGCTCTCTCTCTCTCTCTCTCT 83590  
Qy 7801 CAATAAATTTGCAAAATCATCTAGTCCAAATCTTTTACCAAGTGCCTGAGGCTCCAGAG 7860  
Db 83591 CAATAAATTTGCAAAATCATCTAGTCCAAATCTTTTACCAAGTGCCTGAGGCTCCAGAG 83650  
Qy 7861 AGCATATGCTTCTCAAGGTCACTATAGTGTGGTGAAGTTAGAACCAATCTTAAATTC 7920  
Db 83651 AGCATATGCTTCTCAAGGTCACTATAGTGTGGTGAAGTTAGAACCAATCTTAAATTC 83710  
Qy 7921 CTGACTCCAGTAATCTTGAAGTCAATTTGTTTAAATTTTAAATTTTAAATTTTAAAGAA 7980  
Db 83711 CTGACTCCAGTAATCTTGAAGTCAATTTGTTTAAATTTTAAATTTTAAATTTTAAAGAA 83770  
Qy 7981 TACTTGTAGACACATTAACCCCTGCTGATTAATTAATCTTAAAGTCTCAGGTGATC 8040  
Db 83771 TACTTGTAGACACATTAACCCCTGCTGATTAATTAATCTTAAAGTCTCAGGTGATC 83830  
Qy 8041 C 8041  
Db 83831 C 83831

RESULT 3  
AC031977 19169 bp DNA 1:neat HTG 12-APR-2001  
LOCUS AC031977/c  
DEFINITION Homo sapiens chromosome 1 clone RP11-288018, WORKING DRAFT  
SEQUENCE 3 unordered pieces.  
ACCESSION AC031977

VERSION AC031977.7 GI:13194952  
KEYWORDS HTGS; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULFILL; HTGS\_ACTIVEFIN.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS 1 (bases 1 to 191699)  
Abola, A.P., Bruno, D., Conn, L., Deja Rosa, M., Faulkner, D.,  
Federspiel, N., Glukhov, S., Hansen, N., Herman, Z.S., Hyman, R.,  
Mao, J., Komp, C., Kottler, S., Lam, B., Marathe, R., Miranda, M.,  
Morehouse, A.J., Nguyen, M., Oefner, P., Palm, C.J., Ramirez, D.,  
Southwick, A.M., Webb, C., Wilhelm, J., Yu, S. and Davis, R.W.  
Unpublished  
JOURNAL 2 (bases 1 to 191699)  
Abola, A.P., Bruno, D., Conn, L., Deja Rosa, M., Faulkner, D.,  
Federspiel, N., Glukhov, S., Hansen, N., Herman, Z.S., Hyman, R.,  
Mao, J., Marathe, R., Morehouse, A.J., Oefner, P., Palm, C.J.,  
Ramirez, D., Wilhelm, J., Yu, S. and Davis, R.W.  
Direct Submission  
TITLE Submitted (03-APR-2000) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA  
COMMENT On Mar 4, 2001 this sequence version replaced gi:9665085.  
----- Genome Center  
Center: Stanford DNA Sequencing and Technology Development  
Center  
Center code: SDSYDC  
Web site: http://sequence-www.stanford.edu/group/human/  
Contact: hum-info@sequence.stanford.edu  
----- Project Information  
Center project name: RP11-288018  
Center clone name: 860  
----- Summary Statistics  
Sequencing Vector: M13mp18; X02513; 100% of reads  
Sequencing Vector: plasmid; plasmid\_accession; 0% of reads  
Chemistry: Dye-terminator; 1% of reads  
Chemistry: Dye-terminator Big Dye; 99% of reads  
Assembly program: Phrap; version 0.950319  
Consensus quality: 190680 bases at least Q40  
Consensus quality: 191287 bases at least Q30  
Consensus quality: 191336 bases at least Q20  
Insert size: 195548; agarose-fp  
Insert size: 191499; sum-of-contigs  
Quality coverage: 7.9x in Q20 bases; agarose-fp  
Quality coverage: 8.1x in Q20 bases; sum-of-contigs.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 12646: contig of 12646 bp in length  
\* 12647 12746: gap of unknown length  
\* 12747 94961: contig of 82215 bp in length  
\* 94962 95062: gap of unknown length  
\* 95062 191699: contig of 96638 bp in length.  
FEATURES  
source  
1.191699  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/clone="RP11-288018"  
/clone1db="RPCT human BAC library 11"  
1.12646  
/note="assembly\_name:Contig5"  
12747..94961  
/note="assembly\_name:Contig6  
clone\_end:77"  
95062..191699  
/note="assembly\_name:Contig7

ORIGIN clone\_end:5p6"  
Query Match 99.6%; Score 8009.2; DB 2; Length 191699;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 8028; Conservative 0; Mismatches 13; Indels 1; Gaps 1;  
1 CAATTGAGTCCCTGGGACCTTCAAACTAGAGGCTTAAGTTTACCTTAATTAC 60  
8828 CAAATTAAGTCCAGGACCTTCAAACTAGAGGCTTAAGTTTACCTTAATTAC 8769  
61 AGATTTCCTAGAGATCTAATTAATGGAATCCAGATCGCTCTGTAAAGTTCAAGA 120  
8768 AGATTTCCTAGAGATCTAATTAATGGAATCCAGATCGCTCTGTAAAGTTCAAGA 8709  
121 CTTTCCATGACACCTTACTGTTCTTTCCACCTGACACATGCAATGAACTTATGAAA 180  
8708 CTTTCCATGACACCTTACTGTTCTTTCCACCTGACACATGCAATGAACTTATGAAA 8649  
181 CTGCTGTTTATCTCGGCTAAATGTTGAGAAAAAGATTAACTTTGGATTAAGC 240  
8648 CTGCTGTTTATCTCGGCTAAATGTTGAGAAAAAGATTAACTTTGGATTAAGC 8589  
241 TATTTGGGTTTCTCTTACTCTTGGGAAACAAAGTTTCTTCCCTGCTTAATTAGT 300  
8588 TATTTGGGTTTCTCTTACTCTTGGGAAACAAAGTTTCTTCCCTGCTTAATTAGT 8529  
301 GTGATATGTTCTTCCAGGGAATCAGATGATGATCAGCTGCTATCAATGTCAGG 360  
8528 GTGATATGTTCTTCCAGGGAATCAGATGATGATCAGCTGCTATCAATGTCAGG 8469  
361 TTGAGTTCCTGATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
8468 TTGAGTTCCTGATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 8409  
421 TCCCAAGTGAACATACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
8408 TCCCAAGTGAACATACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8349  
481 AGCCCATATCTTTGCTTTTATGCTTAAATCAATTAATCAATTAATCAATTAAT 540  
8348 AGCCCATATCTTTGCTTTTATGCTTAAATCAATTAATCAATTAATCAATTAAT 8289  
541 GGCATTTGATTAACATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
8288 GGCATTTGATTAACATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8229  
601 ATCTGAATGATGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
8228 ATCTGAATGATGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8169  
661 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
8168 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8109  
721 CTATTTCCGCAATTTGATACGCTCAATCTGAGAAATGTTCTTTGTTGTTCTGCTG 780  
8108 CTATTTCCGCAATTTGATACGCTCAATCTGAGAAATGTTCTTTGTTGTTCTGCTG 8049  
781 AGCAAAGTGGAAAAATTTGAATTTTACTAGAGACACATAGTTCAATCTGCTG 840  
8048 AGCAAAGTGGAAAAATTTGAATTTTACTAGAGACACATAGTTCAATCTGCTG 7989  
841 TGTGGCTGAATGCTGCCCCCGAGTAGAAGAGTCTTTAAGCTTATGTCACAT 900  
7988 TGTGGCTGAATGCTGCCCCCGAGTAGAAGAGTCTTTAAGCTTATGTCACAT 7929  
901 ACCTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
7928 ACCTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7869  
961 CCAATTTCTGATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1020

Db 7868 CCAATTCTCTGATATCCATCTATCTTTTACTGCTCCCAAGCTTTCTTTTAGA 7809  
Qy 1021 TGTAACTGATCGGCTTAAGTACTGAATACCTCTTCTCTCATTACTACAGAGA 1080  
Db 7808 TGTAACTGATCGGCTTAAGTACTGAATACCTCTTCTCTCATTACTACAGAGA 7749  
Qy 1081 AATTCTGATCTTCAAGTACAGAAACACCCTTTAATTCTCGCATTTCTGACCTT 1140  
Db 7748 AATTCTGATCTTCAAGTACAGAAACACCCTTTAATTCTCGCATTTCTGACCTT 7689  
Qy 1141 CATTTATCTCTCTCTCTCTCAAGCTGCAATGAGGGTAAACGCTGATCATTTTC 1200  
Db 7688 CATTTATCTCTCTCTCTCTCAAGCTGCAATGAGGGTAAACGCTGATCATTTTC 7629  
Qy 1201 TCTTTCATCATAGGGCTCGACCAACCACTCTATAGTAAAGACAGTTATAGAGA 1260  
Db 7628 TCTTTCATCATAGGGCTCGACCAACCACTCTATAGTAAAGACAGTTATAGAGA 7569  
Qy 1261 AAACCTAACAAATTTATTTATCAAGTTTATCATGACATGAGAGCTTCAGAAATGAAG 1320  
Db 7568 AAACCTAACAAATTTATTTATCAAGTTTATCATGACATGAGAGCTTCAGAAATGAAG 7509  
Qy 1321 ACCCAAGACCCAGGGGAAACCTGTCTGTCTTTTCTGAGGTTGATGAAGATGATA 1380  
Db 7508 ACCCAAGACCCAGGGGAAACCTGTCTGTCTTTTCTGAGGTTGATGAAGATGATA 7449  
Qy 1381 GCATGTAGCCATGTATGATTAACAAGATATGATGATGTTAAAGACCTCAGGGGGA 1440  
Db 7448 GCATGTAGCCATGTATGATTAACAAGATATGATGATGTTAAAGACCTCAGGGGGA 7389  
Qy 1441 AACACAGAGAGGCTGTATATTCAGATTTCTTCTGATCTCTCTCTATATATAGCAT 1500  
Db 7388 AACACAGAGAGGCTGTATATTCAGATTTCTTCTGATCTCTCTCTATATATAGCAT 7329  
Qy 1501 TCTTCTCTCTGAGTATGAGGACAGACTCTTCTTCAATGAGGCTTTCAAGGAGAGG 1560  
Db 7329 TCTTCTCTCTGAGTATGAGGACAGACTCTTCTTCAATGAGGCTTTCAAGGAGAGG 7269  
Qy 1561 AGAAAGTGGCTTTTATGATTTATGCTGCTCGGAGAGAGAGATTCATTTCTAT 1620  
Db 7268 AGAAAGTGGCTTTTATGATTTATGCTGCTCGGAGAGAGAGATTCATTTCTAT 7209  
Qy 1621 GACCATCTTGGGAGAGAGAAATCTGCTTCTGTGACTCTTCTTCAATGAGAGAG 1680  
Db 7208 GACCATCTTGGGAGAGAGAAATCTGCTTCTGTGACTCTTCTTCAATGAGAGAG 7149  
Qy 1681 GTAAGAGGACGAGGAGGAGATGTCTAGAAAGACCTGCTGCTTCAAGGGCTTCC 1740  
Db 7148 GTAAGAGGACGAGGAGGAGATGTCTAGAAAGACCTGCTGCTTCAAGGGCTTCC 7089  
Qy 1741 GCTCTCTTTAGTTCAGATCTCTTATGACATACCAAGCACTATACCTTGGCATATG 1800  
Db 7088 GCTCTCTTTAGTTCAGATCTCTTATGACATACCAAGCACTATACCTTGGCATATG 7029  
Qy 1801 TTTCTGAGCTCTAACACTGCAATCATGTAACTCTCTATGACCTTCAAACTTCACAT 1860  
Db 7028 TTTCTGAGCTCTAACACTGCAATCATGTAACTCTCTATGACCTTCAAACTTCACAT 6969  
Qy 1861 TGTCTTATTTCTTTATGTTGTGATGAGATGAGTCAATAGCAAGACCTTGAATGCC 1920  
Db 6968 TGTCTTATTTCTTTATGTTGTGATGAGATGAGTCAATAGCAAGACCTTGAATGCC 6909  
Qy 1921 ACTGTCTAGCTGCGATAACATTACTACCACTTAATCAATGTGTAGCTCAGTAAATAC 1980  
Db 6908 ACTGTCTAGCTGCGATAACATTACTACCACTTAATCAATGTGTAGCTCAGTAAATAC 6849  
Qy 1981 TTAAGTCCCTATGCTTATCATCTGTAAATGGAATCAATTAAGAACTTCTTCAAGATG 2040  
Db 6848 TTAAGTCCCTATGCTTATCATCTGTAAATGGAATCAATTAAGAACTTCTTCAAGATG 6789  
Qy 2041 GATCATGAGGATTAAGTGAATTAAGTGCATATAGTCTTAAACCAAGCTTGTCTCAGT 2100  
Db 6788 GATCATGAGGATTAAGTGAATTAAGTGCATATAGTCTTAAACCAAGCTTGTCTCAGT 6729

Qy 2101 TAGTGACAGATAAAATCATCTGTATTAATCTGTGCCCACATTTGTGATGTCTTCTTCT 2160  
Db 6728 TAGTGACAGATAAAATCATCTGTATTAATCTGTGCCCACATTTGTGATGTCTTCTTCT 6669  
Qy 2161 TTGTACAGACATCACTCTATTTATTTATTTTAAAGGCTCTCTGTGAAAACCACTCCA 2220  
Db 6668 TTGTACAGACATCACTCTATTTATTTATTTTAAAGGCTCTCTGTGAAAACCACTCCA 6609  
Qy 2221 GATTCAAAAGATTGAGTTTATCTATCTATCTCTGTGCTTCTCTGAGTTTGTAAAGTAA 2280  
Db 6608 GATTCAAAAGATTGAGTTTATCTATCTATCTCTGTGCTTCTCTGAGTTTGTAAAGTAA 6549  
Qy 2281 ATCTTCACTTGACATCATGATATGTTTGTGAAAATTAACAATTGAGGAGAGACATA 2340  
Db 6548 ATCTTCACTTGACATCATGATATGTTTGTGAAAATTAACAATTGAGGAGAGACATA 6489  
Qy 2341 ACTAAACCAATTTTTTCTCATCAAGTTATATGAAATGGAATGATGAATGAGCA 2400  
Db 6488 ACTAAACCAATTTTTTCTCATCAAGTTATATGAAATGGAATGATGAATGAGCA 6429  
Qy 2401 TTCAAGACCTGCTGTACCTCTTGTTCATCTTAAAGTCACTGTTTCCATATATATGATG 2460  
Db 6428 TTCAAGACCTGCTGTACCTCTTGTTCATCTTAAAGTCACTGTTTCCATATATATGATG 6369  
Qy 2461 ACATTGAGGACTTAT 2520  
Db 6368 ACATTGAGGACTTAT 6309  
Qy 2521 CTGCTAACCTGTCTAACCTGTCTCTCTGCTTCCACACTGTGCCACTCATCTCAGTCTTT 2580  
Db 6308 CTGCTAACCTGTCTAACCTGTCTCTCTGCTTCCACACTGTGCCACTCATCTCAGTCTTT 6249  
Qy 2581 CTTTCAACAAGTACAGATATGATGAGAGACCCCTCTGCTGTCTGTTCTTCCATGAG 2640  
Db 6248 CTTTCAACAAGTACAGATATGATGAGAGACCCCTCTGCTGTCTGTTCTTCCATGAG 6189  
Qy 2641 ATTTCACATGACATCTGATTAAGTCAAGCTCTTCAACAAGCTTCAAAATCCTTGACAG 2700  
Db 6188 ATTTCACATGACATCTGATTAAGTCAAGCTCTTCAACAAGCTTCAAAATCCTTGACAG 6129  
Qy 2701 ATCTATGTTTACTTTCCTTCCATCTCTCTTTATGCTACTTTCATCTGTCTCAATCTCTA 2760  
Db 6128 ATCTATGTTTACTTTCCTTCCATCTCTCTTTATGCTACTTTCATCTGTCTCAATCTCTA 6069  
Qy 2761 GCTATGCTGGCCCTCTCTGCTTCTTCCCA-TTTTTTTTTTAAATTTGTATAT 2819  
Db 6068 GCTATGCTGGCCCTCTCTGCTTCTTCCCA-TTTTTTTTTTAAATTTGTATAT 6009  
Qy 2820 ATTTATGAGTTATAGTGAATCTTTTATGATCATAGTGTATATGTAATTAATCAAG 2879  
Db 6008 ATTTATGAGTTATAGTGAATCTTTTATGATCATAGTGTATATGTAATTAATCAAG 5949  
Qy 2880 GCTTTAGGATTCATCACTGATATGATGATGATGATGATGATGATGATGATGATGATG 2939  
Db 5948 GCTTTAGGATTCATCACTGATATGATGATGATGATGATGATGATGATGATGATGATG 5889  
Qy 2940 ATTCGCTGACTCTTGGCCCTGGATTCATCACTGATGATGATGATGATGATGATGATG 2999  
Db 5888 ATTCGCTGACTCTTGGCCCTGGATTCATCACTGATGATGATGATGATGATGATGATG 5829  
Qy 3000 AAGTAATTTCTCACACACCGGCTGATCTTGGCCCGGATTAATCACTGATGATGATG 3059  
Db 5828 AAGTAATTTCTCACACACCGGCTGATCTTGGCCCGGATTAATCACTGATGATGATG 5769  
Qy 3060 TGCATTTGACCCCTTAAGTATTTCTCACACCGGCTGATCTTGGCCCGGCTGATGATG 3119  
Db 5768 TGCATTTGACCCCTTAAGTATTTCTCACACCGGCTGATCTTGGCCCGGCTGATGATG 5709  
Qy 3120 TGAGGCTGCATTTGCTCATCATCTGACACTATATATATATATATATATATATATATAT 3179  
Db 5708 TGAGGCTGCATTTGCTCATCATCTGACACTATATATATATATATATATATATATATAT 5649

QY	3180	TACTATAAGAGTAAATGCAATAATTTGTCGTCTGTCGTCTGTTTACTATATA	3239
Db	5648	TACTATAAGTAAATGCAATAATTTGTCGTCTGTCGTCTGTTTACTATATA	5589
QY	3240	TAAAGCCCCAGTTCTATAGCGTGTGCAAAAGCATGATTTCTTTTATAG	3299
Db	5588	TAAAGCCCCAGTTCTATAGCGTGTGCAAAAGCATGATTTCTTTTATAG	5529
QY	3300	CTATGTTCTTTTCCCAATTAATTAAGAACCTGCACTTGCTTTACTTATTTGGAA	3359
Db	5528	CTATGTTCTTTTCCCAATTAATTAAGAACCTGCACTTGCTTTACTTATTTGGAA	5469
QY	3360	TACTAATTCAGAGCTTTTGCACTTGCTTTCTCTTCTCAACCATCAATCTATTTAG	3419
Db	5468	TACTAATTCAGAGCTTTTGCACTTGCTTTCTCTTCTCAACCATCAATCTATTTAG	5409
QY	3420	ATACCACTCTTCAAAAGAGGCTTTCTGACCACTTGCTGTAATTAAGCCCTTGACATC	3479
Db	5408	ATACCACTCTTCAAAAGAGGCTTTCTGACCACTTGCTGTAATTAAGCCCTTGACATC	5349
QY	3480	TGATTAAGCTCTAGACATCACTGCCCATTTTATGATGTAAGGTCAAAATCTGGA	3539
Db	5348	TGATTAAGCTCTAGACATCACTGCCCATTTTATGATGTAAGGTCAAAATCTGGA	5289
QY	3540	TCACCTGATTTGTTTATTTTCTGACTCCTTCTACTAGATGAAGAACCTTACTAGACGA	3599
Db	5288	TCACCTGATTTGTTTATTTTCTGACTCCTTCTACTAGATGAAGAACCTTACTAGACGA	5229
QY	3600	GATTTTATCTGCTGTATCAGGTACTGCTTCAAGACACCTGATACAGATAGGTGTC	3659
Db	5228	GATTTTATCTGCTGTATCAGGTACTGCTTCAAGACACCTGATACAGATAGGTGTC	5169
QY	3660	AAAAATATTTCTTAAACAATGAACAATAAAAGTAGACTTTTGAGAGTAAAGCTCT	3719
Db	5168	AAAAATATTTCTTAAACAATGAACAATAAAAGTAGACTTTTGAGAGTAAAGCTCT	5109
QY	3720	TCCACACTACAGAGTCATTTGAGATGACAAATCATAGATTAACGAATTTGATGCTTT	3779
Db	5108	TCCACACTACAGAGTCATTTGAGATGACAAATCATAGATTAACGAATTTGATGCTTT	5049
QY	3780	GTGCATATCAGAGAAAGAGGTGAGAGTGTCAAGTATCATGATGTACAGTCTCTGC	3839
Db	5048	GTGCATATCAGAGAAAGAGGTGAGAGTGTCAAGTATCATGATGTACAGTCTCTGC	4989
QY	3840	CTCTCAAAACAATCTGCAAGTCCACAGTGAAGAAAGTAACTCATGTGAAGCGT	3899
Db	4988	CTCTCAAAACAATCTGCAAGTCCACAGTGAAGAAAGTAACTCATGTGAAGCGT	4929
QY	3900	TTTACAAACATTTTTTAAAGTCTTAAACTCCTAAGAAAGCAATTTATATAGCAAA	3959
Db	4928	TTTACAAACATTTTTTAAAGTCTTAAACTCCTAAGAAAGCAATTTATATAGCAAA	4869
QY	3960	GAAGTGATAAACATGAAATGCTGGAACAGAGTATGAGCTAAGCAACAAGTTAGACA	4019
Db	4868	GAAGTGATAAACATGAAATGCTGGAACAGAGTATGAGCTAAGCAACAAGTTAGACA	4809
QY	4020	TGTTAGTTAAATATGTTCTTGAAGAGAGAGTCTCTGCTTTCAAGAGACAAACAATGG	4079
Db	4808	TGTTAGTTAAATATGTTCTTGAAGAGAGAGTCTCTGCTTTCAAGAGACAAACAATGG	4749
QY	4080	GCAAGTGAACCTCTTGAAATTAATATGTGTAATTAATTTTGGTTATGTCTATAGT	4139
Db	4748	GCAAGTGAACCTCTTGAAATTAATATGTGTAATTAATTTTGGTTATGTCTATAGT	4689
QY	4140	TGTTATATAGATGATTAATAATTTTGAATGACATTTGATGTTTGAATAATCTGAT	4199
Db	4688	TGTTATATAGATGATTAATAATTTTGAATGACATTTGATGTTTGAATAATCTGAT	4629
QY	4200	TTACACAGTTTACCTTATTTGATAGACGTTGAGGATGGGATAGACAGGTGGCAATC	4259
Db	4628	TTACACAGTTTACCTTATTTGATAGACGTTGAGGATGGGATAGACAGGTGGCAATC	4569
QY	4260	CACATTAACGTAGATGAGACACCTGTATCTGCAACCACTGTTAGTTAGAGACTGTA	4319

Db	4568	CAACTAACTGAGTAATCGAGACACCTGATCTGAGCCGAGCTCTGTAAGTAAGCACTGTA	4509
Qy	4320	ACCTCAGCAGACACTCTTCTCTTCTGGGCTCTCATTTCTTTTGGTAAATGAGAGTG	43797
Db	4508	ACCTCAGCAAGTACCTTTCTCTTTCTGGGTCTCATTTCTTTTGGTAAATGAGAGTG	44459
Qy	4380	TTTGGCTAATTTGCCCTTTGGAAGTCCATTTTGCTTTAAGTCCCATCTATTGCGATGT	44339
Db	4448	TTTGGCTAATTTGCCCTTTGGAAGTCCATTTTGCTTTAAGTCCCATCTATTGCGATGT	43899
Qy	4440	TTATATTTAAGTCATGACAAATCAGAGCTCTCTATTCTTAAGTGCAGACATTAACACTT	44599
Db	4388	TTATATTTAAGTCATGACAAATCAGAGCTCTCTATTCTTAAGTGCAGACATTAACACTT	43239
Qy	4500	TATTGTGGAATTTCAAGCATAAGTAATCTTTTGGGACTCATTAAGTTCCGTAATC	45559
Db	4328	TATTGTGGAATTTCAAGCATAAGTAATCTTTTGGGACTCATTAAGTTCCGTAATC	42659
Qy	4560	AATCTATTGAGGATGATCTCTTTTAAAGTGCCAGGTAACAAAGAAAGGCCATGGCTTT	4619
Db	4268	AATCTATTGAGGATGATCTCTTTTAAAGTGCCAGGTAACAAAGAAAGGCCATGGCTTT	4209
Qy	4620	CTTTGAGTACCTTTCTTCCCTTTTAAATTAGTCTGACCTCTTTTAAATGATGAGTCTGACTG	46799
Db	4208	CTTTGAGTACCTTTCTTCCCTTTTAAATTAGTCTGACCTCTTTTAAATGATGAGTCTGACTG	41459
Qy	4680	AATCATTTCCCGTGTCATCTTCTTGATGAGGCTTCCAGTTTCATATTGCACTT	4739
Db	4148	AATCATTTCCCGTGTCATCTTCTTGATGAGGCTTCCAGTTTCATATTGCACTT	4089
Qy	4740	CAGTTCCCTCCACACACACATCAAGATGAGCTGTCAACATTCATTGTTCTATGTTATA	47899
Db	4088	CAGTTCCCTCCACACACACATCAAGATGAGCTGTCAACATTCATTGTTCTATGTTATA	4029
Qy	4800	TTCAAGAAAAGTTGCCAGTAGCTAATTCATTAATGCCCTCTTAATGGGCGGCTAGAGA	4859
Db	4028	TTCAAGAAAAGTTGCCAGTAGCTAATTCATTAATGCCCTCTTAATGGGCGGCTAGAGA	3969
Qy	4860	CTTTTCCCATATTTAATATCATCTCTCTAGATTAATGATGCCCTCCACACATTACATT	4919
Db	3968	CTTTTCCCATATTTAATATCATCTCTCTAGATTAATGATGCCCTCCACACATTACATT	3909
Qy	4920	TGTTGCTGCTCTCTTCTCTCTGCTGCTAGTCATGAAACGTGTTGGTAGTGGGCGAGTGTGGG	4979
Db	3908	TGTTGCTGCTGCTCTTCTCTCTGCTGCTAGTCATGAAACGTGTTGGTAGTGGGCGAGTGTGGG	3849
Qy	4980	AATGTTCAAGGAGCATATGGGTATGGGCGACATATGGGCAATTGCTTTGGGCATCTTTTC	5039
Db	3848	AATGTTCAAGGAGCATATGGGTATGGGCGCGACATATGGGCAATTGCTTTGGGCATCTTTTC	3789
Qy	5040	TATATTTTGGTATTTTGGCATCTCAGTCGAACCCAACTATTTTTCATCTTCCACCTAA	5099
Db	3788	TATATTTTGGTATTTTGGCATCTCAGTCGAACCCAACTATTTTTCATCTTCCACCTAA	3729
Qy	5100	ACATTTTGAATGCCCTGTCTTCTTAATATAATTAAGTCACTGTAGCCATATGATCAGG	5159
Db	3728	ACATTTTGAATGCCCTGTCTTCTTAATATAATTAAGTCACTGTAGCCATATGATCAGG	3669
Qy	5160	AACCTATCGCTTTCTTAATGAAAGCGTGTGTTGGTACATCTAGCAATTAATTCCTCTCT	5219
Db	3668	AACCTATCGCTTTCTTAATGAAAGCGTGTGTTGGTACATCTAGCAATTAATTCCTCTCT	3609
Qy	5220	TCCACTTATAGCTTTCTCTGTATCACTGTGTATGATTTTGGTTATAGGCTATTAAGTG	5279
Db	3608	TCCACTTATAGCTTTCTCTGTATCACTGTGTATGATTTTGGTTATAGGCTATTAAGTG	3549
Qy	5280	TGAAACACCTGAATATCTGTCCATGTCAGGCACTTCAGTTCATATATTGATGTAAAT	5339
Db	3548	TGAAACACCTGAATATCTGTCCATGTCAGGCACTTCAGTTCATATATTGATGTAAAT	3489
Qy	5340	GATPACTGATGTCTAGGTGTTCAAGAAACACCTATAGGGCTTAATATTCCTTAACATCAGTT	5399





```

QY 7560 TCTGACCTTACGCTTAATTTTCTGATGAGTGGAGCAAAATTAATCTTCAAGCTTCACTTA 7619
DB 1268 TCTGACCTTACGCTTAATTTTCTGATGAGTGGAGCAAAATTAATCTTCAAGCTTCACTTA 1209
QY 7620 ACTTGCAATGAAATGAACCGTACAAAATAACAGAGTGTCCAGAGAAAGTTATGCTCTG 7679
DB 1208 ACTTGCAATGAAATGAACCGTACAAAATAACAGAGTGTCCAGAGAAAGTTATGCTCTG 1149
QY 7680 GTAAATATTTTGGAAATGATTAATAAGATTAATCTGAGCTTGTCTCTCAAGAGTTAG 7739
DB 1148 GTAAATATTTTGGAAATGATTAATAAGATTAATCTGAGCTTGTCTCTCAAGAGTTAG 1089
QY 7740 CAGCTAATCTAAGAGAGTAACTCTATGTCCAGAGAGTAACTGCTTCCCTTCCCTCC 7799
DB 1088 CAGCTAATCTAAGAGAGTAACTCTATGTCCAGAGAGTAACTGCTTCCCTTCCCTCC 1029
QY 7800 TCATTAATTTGCAATCATCTAGTCCCAATCTTTACCAACAGTGTCTGAGGCTCCAGAG 7859
DB 1028 TCATTAATTTGCAATCATCTAGTCCCAATCTTTACCAACAGTGTCTGAGGCTCCAGAG 969
QY 7860 GAGCCATGCTCTTCTCAAGGTCAATAGTGTGGTGGTGGTATGAGCAACCAATCTAGAAAT 7919
DB 968 GAGCCATGCTCTTCTCAAGGTCAATAGTGTGGTGGTGGTATGAGCAACCAATCTAGAAAT 909
QY 7920 CCGTACTCCAGTACTCTTGAAGTCAATTTGTTTTTATTTTATGTTTATTTATTAAGA 7979
DB 908 CCGTACTCCAGTACTCTTGAAGTCAATTTGTTTTTATTTTATGTTTATTTATTAAGA 849
QY 7980 ATACTGCTAAGACACTTATCCCTGCAATGATTAATTAACCTAAGATCTCAGGTGGAT 8039
DB 848 ATACTGCTAAGACACTTATCCCTGCAATGATTAATTAACCTAAGATCTCAGGTGGAT 789
QY 8040 CC 8041
DB 788 CC 787

RESULT 4
AC125563/c 275631 bp DNA linear HTG 09-NOV-2002
LOCUS Rattus norvegicus clone CH230-9B12, WORKING DRAFT SEQUENCE, 4
DEFINITION unnumbered pieces.
ACCESSION AC125563.4 GI:24817906
VERSION HTG, HTGS_PHASE1, HTGS_DRAFT, HTGS_FULLTOP.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 275631)
Muzny,D,Marie, Metzker,M,lee, Abramson,S, Adams,C, Alder,J,
Allen,C, Allen,H, Altschrocks,S, Amin,A, Anguiano,D,
Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,
Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
Biswal,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M,
Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E,
Cardenas,V, Carter,K, Cavazos,I, Caesar,H, Center,A,
Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,
Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,
Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D,
Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K,
Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K,
Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G,
Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,
Frazer,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M,
Georgievski,B, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,M,
Gunaratne,P, Haaland,W, Hamill,C, Hamilton,C, Hamilton,K,
Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J,
Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hoques,M,
Hollins,B, Howells,S, Huilyk,S, Hume,J, Idlebird,J, Jackson,A,
Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A,
Karpaty,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C,

```

```

TITLE Unpublished
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Kowis,C, Kraft,C,L, Lebow,H, Levay,J, Lewis,L, Li,Z, Liu,J,
Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,U,
Lorenshewar,L, Louisedge,H, Lozano,R,J, Lu,X, Ma,J,
Maheshwari,M, Mahindraseg,M, Mahmoud,M, Malloy,K, Manum,A,
Mangum,B, Martin,K, Martin,R, Martinez,E,
Mawney,S, McLeod,M,P, McNeill,T,Z, Meenan,E,
Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S,
Morgan,M, Morris,K, Morris,S, Mundasa,M, Murphy,M, Nair,L,
Nankervis,C, Neal,D, Newton,N, Nguyen,D, Norris,S,
Nwachukwu,O, Okwona,G, Olariunpaseon,A, Pal,S, Parks,K,
Pasternak,S, Paul,H, Perez,A, Perez,L, Plankoch,C,
Plopper,F, Poindexter,A, Popovic,D, Primus,E, Pu,L,
Puzo,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M,A, Reigh,R,
Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F,
Rives,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,T,
Sanders,M, Saverly,G, Scherer,S, Scott,G, Shatsman,S, Shen,H,
Shetty,J, Shwartsbeyn,A, Sisson,I, Sitter,C,D, Smay,D,
Sneed,A, Sodergren,E, Song,X,Z, Sotelle,R, Sosa,J,
Steinle,M, Strong,R, Sutton,A, Swalek,A, Taber,P, Taylor,C,
Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Usmani,K,
Valas,R, Vera,V, Villalana,D, Waldron,L, Walker,B, Wang,J,
Wang,O, Wang,S, Warren,J, Warren,R, Wei,X, White,F,
Williams,G, Willson,R, Wleczky,R, Wooden,H, Wortley,K,
Wright,D, Wright,R, Wu,J, Yakub,S, Yen,U, Yoon,L, Yoon,V,
Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Zhao,D, von
Weinstock,G, and Gibbs,R,A, Smith,D,R, Holt,R,A, Smith,H,O,
Direct Submission
Unpublished
2 (bases 1 to 275631)
Wortley,K,C.
Direct Submission
Submitted (29-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 9, 2002 this sequence version replaced gi:23096552.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rac/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GDBK
Center clone name: CH230-9B12
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 228152 bases at least Q40
Consensus quality: 230310 bases at least Q30
Consensus quality: 231872 bases at least Q20
Estimated insert size: 235029; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently

```





QY	5251	GTAGGTAATTTGGTTTATGCTATTAAGATGTGAACCACTGAAATGTTCTGTCCATGCAAG	5310
Db	143600	CCCTTTGTATGTGATATCTGGGATATTTTAAACCAGATCAACGATGTGATTAATAGGCG	143544
QY	5311	CATTTCAGTTCATGATATTTGTATGTAAAGATACATGATGT- - - - -CTAGGTGTC	5361
Db	143540	CTTTCAAGAACCCGGAAGAAATTCAGCCCATGATACGTAAATGTGGAAGATTCACAAACTC	143481
QY	5362	AGAAACCTTATGCGGCTTAATATTCTTCA- - - - -ATCAGTTGAAGCTGTGTATCG	5416
Db	143480	AGAAAGCAACGACGGGACGGCTCTCCACCGGTCTGTCCCTCAACCTGCTGCAACAG	143421
QY	5417	CAAGCAAACTACATATATTTTCTGCTGCTCTCT- - - - -CTCTTCTCTCATCTCTCTT	5473
Db	143420	CAAGCGAATGACATCTTCCGTGCTGTGATCTCTCTCTCTCTCTCTCTCTCTCTA	143361
QY	5474	TCTTTATCTTTTGAATATCAGTTTGTGACATCTTAAGATTAATCAATGAATAAACCAATT	5533
Db	143360	GTTTTCTCTTTTGAACAATTAAGATTCACGTTACA- - - - -GACATACAGTAAATCTGACT	143303
QY	5534	GATATAAGAAATGCTGTGTATATTTGCTCATCTACTCCTCCTTGTCCTCGAGCTGCC	5593
Db	143302	GAGAT- - - - -GGAATGTGTGTCT- - - - -TATGTATTATCAACCCCTGTGTTGGT- - - - -CTTAAGTTGCT	143246
QY	5594	GGTTTAACTTTTATACAGACGACGACATGTGAAGAGAAAC- - - - -TGTCAGTGTAGGTTGA	5652
Db	143245	GTTTCACACTTTTGACAGAAATCGAGCTTCTTGAAGAAAGACCCGACACTGTGGCTGA	143186
QY	5653	ATTCTGTGTATACCAAGATTTCTAGAAAAGTATTCCTCAGTCAGGTTGATTAACAGATATA	5712
Db	143185	TTTTTGTACTAGTAACGAGATTAGGA- - - - -TGTGCAAAAGTAATATGATCAATCATATA	143128
QY	5713	GCAATATATTTTTTCTAGGTAATTTCTGTATGTCTGCCGGCTTATATACGTCTGTTCAT	5772
Db	143127	CTGATTTAAATTCCTCTGGCTGTGTTTCCCTACGCTGCCGACATATAGCTGTCTGAT	143068
QY	5773	CCAAGTAATTTCTCCACCTTCTGTGTTGCAATACCAACGACAACTTCCGCAATCAC	5832
Db	143067	CCAATCCCTGTGCTTATTTTATTCG- - - - -CAGGCTTATTTTATCAATCAC	143017
QY	5833	TGCGTGAGACGATATTCCTGCGAGCTCCCTTTTGGAATGTGTAGATTCAGATCTTGA	5892
Db	143016	TGCCCAAGACCTCGATTTAGTACCAACTCTGCTTA- - - - -	142979
QY	5893	CCATGTATAATATGATGCTTCTATTCAAAAGAGAAAGGATGGAGTGCAGCTCTTA	5952
Db	142978	- - - - -CACACCCAGATATATTTCTGACTCAAGAAAGAAATGTATGTAGAGCCAGCGCAA	142920
QY	5953	GTAAGCCCGAATTCCTGCTGTGTAATTTTCTCTCAGAAAGCAACTCTGTGATATTT	6012
Db	142919	ATAAGATTAAGATCCCTGTGTGTTCTTTTTCGTGAGGCAAGTT- - - - -CTTTAGACAT	142861
QY	6013	TTTTTTACAGSCATATGAATAAAATAATATTTTTCAG- - - - -CATTTACACTTTTTCCT	6070
Db	142860	GATTTTACAGCCCAAGAAATAGTCTCATATTTTTCAGAACCAATGATTCATATGATCT	142801
QY	6071	TTTTCTGAATTTTAACCTCTGACATGTGTGTGACATCTAGTAAATCTTTTCCCAT	6130
Db	142800	CTCTTACAAATTTAGAACTTACACATGTGTGAAATGTAGCAA- - - - -ACGTCC	142748
QY	6131	CCCTACTTTTCAGAAAGATTTTCTGCTGTTCACTTAACTTGATGCTGACTCT	6190
Db	142747	TTTTCTGTTCGAAAGCTTCTCCCTCGCTGGGTCTCTTTGATCCCCCATGCGTCAAGTAG	142688
QY	6191	TTTTCTTCTCATCTCTTTCAGGGGGCTGAGAGAGCAGAGGACAGAG- - - - -GAGCTGGTA	6247
Db	142687	TTTTCTTCCC- - - - -CTTTCTAAAGGCTTTAGAGGCAAGAGGCTCTACAGCAGAACGATC	142630
QY	6248	CTGCAAGAGCGGTGCTGTATGTGCTGAGCGGT- - - - -GTAAGTGGGCTATTAAGAGACCCCT	6306
Db	142629	CTATAGCCCTGCATGTATGTGCTGCGCGGCGCTGTGCTGAGCTATTAAGAGACCCCT	142570

OY	6307	ACAGGCTTAGCAGGAAGACCTCACAAGATTCTGCAAATATCTTTACCAGGAAGAAGCA	6366
Db	142569	ACAGGCTTAGAGGCGAAGGCTCAAGAGATTCGTACAGCTTCTTTGCAGAGCAGAAA--C	142512
OY	6367	AAGAACGCTCAAAGCCGAAGCACACAGCTCTCTCCGCCATTTCTTTCTGCTTGCCAAT	6426
Db	142511	AACGTGCTCAAAATAGAAACCAACGGTGTCTCCGCTGCTTCTCTCTCCGCCAAT	1424525
OY	6427	TCCAAGCTGTTAAATAGATGTGC AAAAGGCGCTTGAGGCTCTCCGCGCTTCTTGAGG	6486
Db	142451	CCCA--CAGTTAAACAAGATGTGC AAAGAGACTCGGTGTCTGCGCGCTTCTGCTGAGG	142394
OY	6487	AGGAAGATTGCTTTAGCATTAACCAATATAAAGCTTTGGCTAGACTTCTCAGTTAT	6546
Db	142393	AGGTAAGATTGCGTTCTGTGTACTGACCAAACTAAAGTTTGC AAGACCTTCTCAAGAG	142334
OY	6547	TTACATGTTGTACTCTACCTAGTCTGTGCATTAGAAACAGTGTGCAGAGAGC	6606
Db	142333	TTGAGGCTTGCTGTATGAAC---TTCTGTGGATCCGAGGCTATAGGCTCTGGGAAGC	142278
OY	6607	ACGACTTTCTAACTTTCTCTC CAAGCTAGTATGATTTGGACTTAAGACATGTGCTCCC	6666
Db	142277	CC----TGGTATGTGCCCCAGACCTAGCAGACAGTACTTCGAAAGCCGGCTCT	142223
OY	6667	CAAAATTCAGCCCTTATGTGTGG--TTTTGTGTGACCTCAGTTTGAACGTGTTCTATT	6724
Db	142222	CAAAATGTGCTCTGTGTGTGTGTGTTTGATTTGGATGCTTTAGTGTGGGACTGGTCTGAC	142163
OY	6725	CTTTAAACCAAGCTC--TAGAAAGCTGTATTAAATTAAGAGCGAATGAGTTGAGG	6781
Db	142162	TTTTAAACCTGGGCTCTTGTATTAGCTGCTGGATTACGAAGGACAGATGAGTTGGGG	142103
OY	6782	CTATGTACAGTATCTGTAAATCTCCACATGTGTATTCTACTCTATTGAGATCCCT	6841
Db	142102	-----GATGATCTGAATTCCTTCTGTGATTATTTGCTCCTGTTTGACCATTTGCT	142054
OY	6842	GAGATCATAAGAACCTGGGCTCTGGGCTTTCTGATTGTATGCTACAACTTGTTCAGAA	6901
Db	142053	G-----AGGACCAAGAGCAGGGCTTAATGATT-----CTTGTGGAGGA	142013
OY	6902	AGGTAACCCAGATAGAGTTGGCTCCATCAATCAGAAAGACATAGCTTCCGTGGGT	6961
Db	142012	AGAAACCTCAGAAAGAGTTTGGAAACCTGCTCGAAA-----CCTTTCTGCACAT	141961
OY	6962	GGTCAAGTAATCTCTCTCTATGTTCTTTAAGCAAAATTTCAATGATGATGAGTT	7021
Db	141960	TCTTCACTGTCTCTACTCTAAT-TAGTTGAAAAAGAGCTGAAATAAAGTTCTTT	141902
OY	7022	TTTAAAGCAGATTCCTTATCTCTCTGCCCCATCTTAGTCTTGAAGTGTCTCATATG	7081
Db	141901	TCCAAAGTCAATCTTCTTATCTCCCTCCCGCTCCAGATGTGTGAAGGAGTGTGGT	141842
OY	7082	AGTTGTGAGAAATATGATCATTACAAATCAGTTAATAGTTTGTAGAGATCTCAT	7141
Db	141841	TGAGAAATTAATATACAGTTTGTACCTACCMNNNNNNNNNNNNNNNNNNNNNNNNNNN	141782
OY	7142	CTTAAAGACATGTTTGTGTATATATCTCCCTGATTTTAAAGACCTTACAGACAT	7201
Db	141781	NN	141722
OY	7202	ACAGCTATTCATTGTTTTGTGTTGTTCTCAAAAAGATATAAGAAATGATTCAGAGAA	7261
Db	141721	NN	141662
OY	7262	AGATCATATATTAGCAGTTGAAAAATTAACAACAAAATGAGTGACATTTACATTATPA	7321
Db	141661	NN	141602
OY	7322	TCTTGACATCAAAAGTAAAAAGTCAACTAAAGTATACACCTGCTTCTTATGCACCT	7381
Db	141601	NN	141542
OY	7382	GCAAAATAGAAATTAACAATAATTTATTTGGAATATATCTCAGAAACAATAATTTTATA	7441

[illegible]

**TITLE**  
**JOURNAL**  
**COMMENT**

Submitted (13-Dec-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Dec 13, 2003 this sequence version replaced gt:39540543.  
 All repeats were identified using RepeatMasker:  
 Smith, A.F.A. & Green, P. (1998-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center -----  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIRB  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [submissions@genome.wi.mit.edu](mailto:submissions@genome.wi.mit.edu)  
 ----- Project Information -----  
 Center project name: L23317  
 Center clone name: 82\_I\_24

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 9 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

1      71982: contig of 71982 bp in length
*      71983      72082: gap of 100 bp
*      72083      117517: contig of 45435 bp in length
*      117518      117617: gap of 100 bp
*      117618      122151: contig of 4534 bp in length
*      122152      122251: gap of 100 bp
*      122252      126124: contig of 3873 bp in length
*      126125      126224: gap of 100 bp
*      126225      163226: contig of 37002 bp in length
*      163227      165826: gap of 100 bp
*      165827      165926: gap of 100 bp
*      165927      169052: contig of 3126 bp in length
*      169053      169152: gap of 100 bp
*      169153      175011: contig of 5859 bp in length
*      175012      175111: gap of 100 bp
*      175112      224486: contig of 49375 bp in length.
  
```

**FEATURES**  
 source  
 1..224486  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="1"  
 /map="1"  
 /clone="RP23-82124"  
 /clone\_lib="RPCT-23 Female Mouse BAC"

**ORIGIN**  
 Query Match 5.0%; Score 402.6; DB 2; Length 224486;  
 Best Local Similarity 54.1%; Pred. No. 1.7e-73;  
 Matches 1728; Conservative 0; Mismatches 1174; Indels 301; Gaps 33;

```

OY      3561  TGACCTCTTCTACTGAGATGAAAACCTACTACTAGAGCGAGATTTAATCTGCTGTATCAG 3620
Db      128421 TGCGTCCCTGTTCTCCGTCCCTCCCTGCAAACTACCTCCCATCTCACTTTTGGTATGTA 128480
OY      3621  GAACTGCTTCAACAGACGACCTGATACAGAGAGGTCGCAAAAGATATTTCTTAACAAA 3680
Db      128481 GCTTAAGTAAATCTGTGCTATGATTTACCTGTGTGAGAGATGCTCAGGTGTGTGAAAG 128540
OY      3681  TGAACAAATAAAAAGTAAATCTTTTGAAGAGTAAAGCTCTTCCACACTACCAAGATCATTG 3740
Db      128541 TGAGCTTGACAGAGAGACCTTTGAGAGCAATCCCATTTAC--ATAGAGCAAGC 128597
OY      3741  AGGAATGCAAAATCTTAAATATTAACGAATTGATGCTTGTGCAATATACAGAAAGAAAG 3800
Db      128598 AGGAATGCAAAATCTTAAATATTAACGAATTGATGCTTGTGCAATATACAGAGAGGAGAG 128657
  
```

OY	3801	TGAAAGTGTGCAAGGATACAGATGACAGTCCCTGCCCTCTCAACA-----CAATC	3855
Dp	128658	TGAGCCATTGTCAAGGGTATCTCGATGTACAGTCTCGCTCTTCAACAAACAGCCAGT	128717
OY	3856	TGCAATGCCACAGTGAATAAGTTAACTCATGTGAAGCGTTTTCAAACACTTTT	3915
Dp	128718	TTGATGTAAAAATAAAAATAAAAAGCTAATTTATGAAGAACCTTTTCAATA---TTGG	128774
OY	3916	TAAAGTCTTAAACTCCTTAAGAAAGCAAGATTTTATGTCAAGAAAGTGAATAAC---	3972
Dp	128775	TGGAGATATTAGACTGTGGGAATGCAAGGACAGTCACTAAGTGAAGATATGACTTG	128834
OY	3973	-----ATGAAATGCTGAACAGATATAGCTTAAGCCAAAGTTAGACATGTTA	4024
Dp	128835	GGCAGATGAAGAGTTGAAGAACTGGTTTTCTTGGGAAGCAATCAGGCAATCTGCTATAG	128894
OY	4025	GTTAAATATGCTTGAAGAGCAGCAGCTCCGCTTTCAAGACCAAGA-----	4070
Dp	128895	CTCCTGTGTCCTGGAACAAGAGAAATCATGCAAGAGAGAGAGAGCCAGAGAGA	128954
OY	4071	-----ACAAATGGCGCAAGTGAACCTCTTGAAATAAATGTGTAATAATTTT	4121
Dp	128955	CAGAGAGACAGAGATCTCACAGAAATCTTGCGCAAAATGTAGAAATTAATTAATTT	129014
OY	4122	GGGTTATGTCTATCTACTGTGTAAATAGATATATAAATATTTGACTGACACTTGTGA	4181
Dp	129015	GACCTTTATATCTGCTATCTATGTTGTGACATATAAATATTTGGCAAGCACTTGCA	129074
OY	4182	GTTTAGAATATCTCTATTTTACACAGTTTACCTTATTTGATTAAGCTGTTGAGTATGG	4241
Dp	129075	GTTTAAAAATCTCTATGTATACACAGCATATCTCTTGATTAAGAAACATTAAGATAGA	129134
OY	4242	ATAGCATGCGGACATCACATTAATCTGATGTGAGACACCTTATCTGGACCCAGCTC	4301
Dp	129135	ACAGCTCTGGGACAGCTACACAAACCAAGTGTCAAGAACTGGTCTGATCTGTGCG	129194
OY	4302	TGTTAGTAAAGACTGTAA---CTCAGCAAGTCACTTCTCTTCTGGGCTCTATTTTC	4358
Dp	129195	GGTTAACAGAAATTTGTATCTTGGAGGCGTGCCTTTCAGTCTCTGGGATTCCTATTTA	129254
OY	4359	CTTTTGTGGAATGAGATGTATAGCTAGATGCGTTTGAAGTCCCATTTTGTCTTAA	4418
Dp	129255	TCTTGGGTGAATAGGGGTACAGATAG---TGGTTTTAA	129293
OY	4419	AGTCCCATCTATTCAGTATTTATTTAACTATACAAATCAGCGTCTCTTATTTCT	4478
Dp	129294	AGGTCAATTAGCTCAGTATAGCTAATGC-----TATACCTCTTCTCTTAGACCG	129347
OY	4479	AAGTGCACAACATAAACTTTATTTGTGGAATTTGAGCATCAGTAATCTTTTGGGTA	4538
Dp	129348	AGATAAAAAGTCTCTCTCGGAGTTCTCTTTTCAACACAGGATCATTTCTTTGAGTC	129407
OY	4539	CTC-ACCTTATGTTCTCGAATTAATCTATTTGAGTATCATCTTTTACGTGCCAGGTA	4597
Dp	129408	TTCTATGTATGTTCTTAAACCAAGATGTCTGAGTCCCTCGTGGGCTCAGGTA	129467
OY	4598	AACAAAGAGGCGATGCTTCTTCTTGAGTCACTCTTCTCCCTTTAATTAGTACCC	4657
Dp	129468	AGGAAGAACCTCTAGTCACTGTGGGAAACAGCGCTCTCTTTTAGTTGATCT----	129523
OY	4658	TCTTTATGTCACTTTGACTGATTCATTTCCCTGTCATCTTCTTGCTGAGGGCC	4717
Dp	129524	TCTTTATGCAATTTCTGCTG-----CAATTCAGTCCAGGTC	129583
OY	4718	TTCTTAGTTCAATATGCACTCAGTTCCTTCCACACACCATTAAGATGGCTGTAAAC	4777
Dp	129564	TCTCTCACTTCACTCGGACCTGATCTCCTCCGATCCACACCAAGAGACGCTCATTC	129623
OY	4778	ATTTC-----ATTGTCTATAGTTATATTCAGAAAGTTGCCAGTAGCTAATCCAAAT	4833
Dp	129624	TTTCTATCTAGGTTTCACTGTATGAAAGCAAGAGAGGTTTTTC--ATAGCTGAATTAATA	129681

QY	4834	AAAGCCCTCTTATGGCGGGCTAGAGACTTTTCTCTAATTAATTAATGACATCTCTGAGA	4893
Db	129682	AGATTTTGTGTGTTTTTTGTTTTGTTTTGTTTTTATTATTTTCTTTACAGAGAGAG	129741
QY	4894	TTATGTCCTCTCCACCACTTTACAT-----TTGTCTGTCTCTCTGCTCTGTATCA	4948
Db	129742	CTGGGGAATCTTCTCCATACACTGACATACCTATATGCCCTTTTATGTTTTGTCTGTATCTC	129801
QY	4949	TGGACGTTGTGTAGTGGGGGCACTGTGGGATGTTCAAGGGACGATTTGGTATGGCC	5008
Db	129802	CATATCACTACCTGACCAAGGAAAGATGTATGACATTAAGAGGGCAGGTGTTGTCCAGAGC	129861
QY	5009	ACATATGGGCATTTGCTTTGTGCATTT-----CTTCTATATTTTGGTA-----	5052
Db	129862	ATAGAAAGAACATTAACATGCTCTTATTTCCCACTTAATTTTGTATTCCTCTAGAGAAA	129921
QY	5053	-----TTTTGCATCTCACTGGAACCAA	5075
Db	129922	CCCTCTCAAAAGAAAGCATGTGACCTCTCTTTCTTGTGCAAAATCACCAGATGCT	129981
QY	5076	CTATTTTTTATCTCTCCACCTTAACATTTTGAATGCTCTGTTC-TTATATTAAGA	5134
Db	129982	GTTTTTATTTACTTTCAAGCCCAAAAGACTGAAAGACGTATGCTTTGTGTGAATAATC	130041
QY	5135	TAGCTCACTGATGCTATGATCAGAACCTATCTGCTTTCATATGAAGACGTGTTGGT	5194
Db	130042	TAGGTCACTG--GATTAAGACCAAGCACTGTGTGACTTTAAGACGAATCTGTTTGGT	130059
QY	5195	CAGATCTACCAATTAATTCCTCTTCCACTTAATAGCTTTCCTCTGTA-----ACT	5245
Db	130100	CACACACTAACAACCAACCTCTCTCTTCCAGTACACATCCCCAGTGTACACGCCAAT	130159
QY	5246	CTGTGTAGGTATTTGTGTTATGTGCTATAGATGTGAACAACCTGAATGATTCGTCCAT	5305
Db	130160	TTGCTGTGCTATCTGGGATTTCAACCATGATGTGAAGATGTGAATCTTCAGAGAT	130219
QY	5306	GCAGGATTTCACTTCATGATATTTGTATGTAAAGATACGATTTGTACGTGTCAAA	5365
Db	130220	CTGAAGATTTCACTGCGTGGCAGTAATGTAAAGATTTCT-----TGAATCTCAAG	130272
QY	5366	ACACCTAATAGGGCTTAATATTTCTTAACAATCAGTTTGAAGCTGTGTATAGCAAAACAA	5425
Db	130273	GCCACAGCAGGATGGCTGCTCTCCACCGTTTCAAGCCTGTATATGCAAGGCAAA	130332
QY	5426	CTACATATTTTCTGCGCTGC-----TCTCTCTCTTCTCTCAATCTCTCTTCTTT	5478
Db	130333	CTACACATCTTTCTGTGCTGTGATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	130392
QY	5479	ATCTTTTGAATATC-----AGTTTG	5499
Db	130393	CTCTCTCTCTCTCCCCCTCCCTCCCTCCCTCCCTCCCTCTCTCTTTTGAATAAATA	130452
QY	5500	GACACTTAATTAATCATAGACATTAACCACTTGATATAAGATTGCTGTATATTTG	5559
Db	130453	AAATTTACATTTACATATCATACATCAATCTCGGTCGACATTGGAAATGTGTTTAATG	130512
QY	5560	CTCATCTACTCCCTCCTTTGTGTCTGTGACGCTGCGGTTTAGACTTTTACAGACGACG	5619
Db	130513	TTCATCCACCCCTGCTTGTGTCTTTTAGTC-----TTTGCAGAGTCAAG	130559
QY	5620	CATGTGAAGAAACATCTCACTAGTGTAGCTGATTTCTGTATTACCAATTTCTAGAA	5679
Db	130560	GGTCTTCAGAGAAAGAACGCAAGCTGACATGATTTTGTGCTCAGTAATATTATAGGA	130619
QY	5680	AAATATTCCTCACTGCTGATTAATTAACATATAGCAATATCATTTTCTAGAGGTAGTT	5739
Db	130620	ATGTCT--CAAGTCAATATGATTTACAGACATCTGATTTAATCTTCAT-----TGGCTT	130673
QY	5740	CTGTATGCTGCGGGGCTTATAACTGTCTGTATCCAGCTATTTCTCCACCTCTTGTT	5799
Db	130674	TGCTATGCTGCGGCAATATAGCTGTTGTGCTATCA-----CATCCTCCGTGT	130722
QY	5800	TGCATTAACCAACAGCACTTCGCAATATCATGCGTGTGAGACGATGATCTTGCAAGCT	5859

Db	130723	CPACCAATCAGGCGACGATATTTTCATTAATACACTGCGCCAGAGACCAAGATCTTAGTACCA	130782
Qy	5860	CCCTTTTGGAAATCGTGAGATCAGATCTTGGACCATCTATTAATATGATGCTTTATTC	5919
Db	130783	GGCTGECTCA-----CACTACCAAAATATATTTCTAACCC	130819
Qy	5920	AAAAGAGAAAGGACATTTGGAGATCAGCTCCTAAGTAAGCTCCAAATTTCTCGTGGTACT	5979
Db	130820	AAGGAAACATCTATTTAGAACCCAGAGTGAAGTGAATGAACATCTTGTGGTTTC	130879
Qy	5980	TTTCTCTCCAGAAACACTTCTTGATATTTTTTTTACAGGCATATGATTAATAACT	6039
Db	130880	TTTCTTTCTGGAGAAACAGTTTCTTAGCATTTGATCTTACAAAGC-----AATGAAT	130931
Qy	6040	AATTTTGGAGATATGACACTTTTTCCTTTTCTAGAAATTCATAACCTCTGACATTG	6099
Db	130932	ATTTTTCAGCAGCGTGATTTCTATTTTCTTTTCTTCTTAATAATGAAACTTTGACGCTAG	130991
Qy	6100	GTGGACACATTGAG--TACATTTTTTCCCATTCCTACTTTTCAAGAGATTTTCTCT	6156
Db	130992	GTGAAACGTTAAGCAAACCTCATTTTCTATATCCTACTTTGCAATGGCTTCTCTCC	131051
Qy	6157	GCTCTTCACTTAACATTTGCTATATGCGTCAAGCTTTTCTTCTCATCTTTTACGAGGCT	6216
Db	131052	GCTGGGTCATTTGATCCCTGTATGCGTATGAGTGTGTTTCTTCTCACCTTTCTGAGGGCT	131111
Qy	6217	GGAGAGGCGAGAGGAGACAGAGAGACTGGTACTGCAGACGCGTCTGTATGGCTGAC	6276
Db	131112	GGAGAGGACGCGCGCGCGAGAGAGCGGAGCGCTGAGACCTGTATCTGATTTGGCTGCG	131171
Qy	6277	GGT-GGTAGCTGGGCTATTAAGAAGACCCCTTACAGGCTTACAGAGAGAGCGCTAGAGGA	6335
Db	131172	GGGCTGGCTAGACTATMAAGAACCCCAACCGGGCTTGAAGCGCAGAGGCTCAGAGGA	131231
Qy	6336	TTCTGCATATCTTTTACCGGAGAGAGGCAAGAGCTCAAGCGGAGCCACAGCTC	6395
Db	131232	TTCTGCACACTTCTTTTGCAGACGAAG--CACTGTCTCAAAATAGAAACACCGCGG	131285
Qy	6396	CTCTGCGGCATTTCTTTCTCTGCTTGCAGATTCAGCGTTTAATAATGATGCAAAAG	6455
Db	131290	CTCCGCTCTCTTTCTCTCTCGCTAAGATCCC--TCAGTTAAACAAGATGTCAAAAG	131344
Qy	6456	GCTTGAGGTCGCGCGCTTCTTGTCTTAGAGAGGTAATTTGCTTACGCATTAAACAT	6515
Db	131348	ACTTGAGGATCTGCGCGCTCTCTGCTAGGAGGTAATTTGTTCTGTTACTGCTCAG	131407
Qy	6516	ATTAACTTTTGGCTAGACTTTCTCAGTATTTTACATTTTGACTTACTAATCAGTTCT	6575
Db	131408	GCTAAAGTTTTCATAGACTTTTTCAGTAGAGTTGATGCT---TTTGCTTAGAAAGTTCT	131466
Qy	6576	GTGCAATTGAAACAGCTGTGTTCAGAGAGACAC	6608
Db	131465	CTGGAGTCAGAAAGCTATATGCTCCAGAGAGCCC	131497
RESULT 6			
LOCUS	AC006581	112931 bp	DNA linear PRI 22-JUN-2002
DEFINITION	Homo sapiens 12 BAC RP11-259018 (Roswell Park Cancer Institute		
ACCESSION	AC006581		
VERSION	AC006581.16	GI:4914350	
KEYWORDS	HTG.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 172931)		
	Muzny, D., Aronson, A.P., Bouck, J., Bunac, C., Chen, Z., Ding, Y., Dugan, S., Durbin, J., Fortnum, J., Garcia, C., Gorrell, J.H., Gorrell, L.L., Hernandez, J., Issar, A., Jackson, L., Kneitz, S., Kondejewski, N., Lau, S., Leal, B., Lee, E., Lichtege, O., Liu, W.,		

TITLE	JOURNAL
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 172931)
TITLE	Worley,K.C.
JOURNAL	Submitted Submission
REFERENCE	Submitted (19-FEB-1999) Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
AUTHORS	3 (bases 1 to 172931)
TITLE	Worley,K.C.
JOURNAL	Direct Submission
REFERENCE	Submitted (28-MAY-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
AUTHORS	4 (bases 1 to 172931)
TITLE	Worley,K.C.
JOURNAL	Direct Submission
REFERENCE	Submitted (03-JUN-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston TX 77030, USA
AUTHORS	5 (bases 1 to 172931)
TITLE	Worley,K.C.
JOURNAL	Direct Submission
REFERENCE	Submitted (16-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston TX 77030, USA
AUTHORS	6 (bases 1 to 172931)
TITLE	Worley,K.C.
JOURNAL	Direct Submission
REFERENCE	Submitted (22-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
AUTHORS	On May 28, 1999 this sequence version replaced gi:4844002.
TITLE	INFORMATION: <a href="http://www.hgsc.bcm.tmc.edu/">http://www.hgsc.bcm.tmc.edu/</a> or email <a href="mailto:gc-help@bcm.tmc.edu">gc-help@bcm.tmc.edu</a>
JOURNAL	

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Gene and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarly (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: [http://gc.bcm.tmc.edu:8086/quality\\_info/genbank\\_annotation.html](http://gc.bcm.tmc.edu:8086/quality_info/genbank_annotation.html).



Db 48503 CTTGTTTGGGAGAGAGTTCTAGTTCTTACCACTTGGGGAAGAGATTATAG 48562

Qy 1648 GTTTCGTGACTGCTTCATGAGAAAGAGAGTAAGAGCAGAGGAGAGATGCT 1707

Db 48563 GTTTCGTGACT--CCTAGGCGAGAAAGAGAGGTGAGACAGAGGCGAGAGAGCT 48620

Qy 1708 CAGAAAGAGACTGGCTCTTCTGAGGGCTTCGCTCTCCTTAGTTCGAAGTACTTCTT 1767

Db 48621 CAGAGA-----CTTTCCTCTGAGG-----CTTTTAGTCTCTAAGTACTTC 48662

Qy 1768 AGCATACCAAGCACTACTCTTGGCATATGCTTTTGGCTCTAAGTCTGAA 1822

Db 48663 AGCATGCCAAGGCCCACTCTTGGGTAATCTTTCTGAGCTTCACCAAGCTA 48717

RESULT 7  
AC119619/c

LOCUS AC119619.5 94158 bp DNA linear PRI 28-SEP-2002

DEFINITION BAC Library) complete sequence.

AC119619 AC012577

AC119619.5 GI:23343650

VERSION HTG.

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

1 (bases 1 to 94158)

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alstrooks, S.L., Amaratunga, H.C., Are, J.R., Banks, T., Barbarella, J., Benton, J., Bimaga, K., Blankenburg, K., Bonini, D., Bouck, J., Bowe, S., Brieve, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Dem, A.L., Ding, Y., Dinh, H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Barnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escoto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., France, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolyet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korval, J., Kovar, C., Kravovic, J., Kuresh, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lien, C., Liu, J., Liu, W., Louissegh, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mel, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, J., Nwokwenkwo, S., Ogutu, M., Okwomodu, G., Oragunye, N., Oyedero, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Sher, H., Shooshari, N., Sisson, I., Sodergren, E., Sonake, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telirod, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vanson, R., Wang, C., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G. and Gibbs, R.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 94158)

AUTHORS Worley, K.C.

TITLE Direct Submission

JOURNAL Submitted (29-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE

AUTHORS Worley, K.C.

TITLE Direct Submission

JOURNAL Submitted (19-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 94158)

Worley, K.C.

REFERENCE

AUTHORS Worley, K.C.

TITLE Direct Submission

JOURNAL Submitted (28-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Sep 28, 2002 this sequence version replaced gi:22297319.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

COMMENT

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht>

ml.

FEATURES

source

Location/Qualifiers

1..94158

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="X"

/clone="RP11-1J4"

complement(1..2004)

/note="overlaps bases 1..2004 of clone AC112499"

/function="clone overlap"

4..92

/rpt\_family="L1MB2"

complement(479..781)

/rpt\_family="MER58B"

1591..1925

/standard\_name="SHGC-111221"

1635..1925

/standard\_name="SHGC-68548"

complement(1803..1861)

/rpt\_family="MER91A"

repeat\_region

STS

STS

repeat\_region

```
repeat_region complement(2158..2465)
/rpt_family="AluSx"
repeat_region 3075..3245
/rpt_family="MIR"
repeat_region complement(3246..3436)
/rpt_family="LTR54"
repeat_region 3437..3739
/rpt_family="AluO"
repeat_region complement(3740..4007)
/rpt_family="LTR54"
repeat_region 4008..4046
/rpt_family="MIR"
repeat_region complement(4405..4640)
/rpt_family="MER30"
repeat_region 4834..5153
/rpt_family="MER4D"
repeat_region complement(5151..5592)
/rpt_family="LIM4"
repeat_region complement(7670..7719)
/rpt_family="L2"
repeat_region complement(7872..8182)
/rpt_family="AluSg"
repeat_region complement(8525..9084)
/rpt_family="MLTIE"
repeat_region complement(9362..9650)
/rpt_family="MLT1A1"
repeat_region 10392..10501
/rpt_family="TGA)n"
repeat_region complement(11028..11174)
/rpt_family="MER5A"
repeat_region 11717..11808
/rpt_family="MIR"
repeat_region 11937..12248
/rpt_family="Char1ie4a"
repeat_region 12249..12463
/rpt_family="MER20"
repeat_region 12464..12610
/rpt_family="Char1ie4a"
repeat_region 12611..12947
/rpt_family="THE1B"
repeat_region 14522..14913
/rpt_family="THE1B"
repeat_region 14914..14947
/rpt_family="Char1ie4a"
repeat_region 15035..15204
/rpt_family="MIR"
repeat_region 15205..15423
/rpt_family="MER58A"
repeat_region 15424..15519
/rpt_family="MIR"
repeat_region 15620..15661
/rpt_family="(TA)n"
repeat_region complement(16448..16619)
/rpt_family="MER968"
repeat_region 16639..16659
/rpt_family="AT_rich"
repeat_region complement(17007..17274)
/rpt_family="AluJb"
repeat_region 17441..19201
/rpt_family="L1PAS"
repeat_region complement(19202..19496)
/rpt_family="AluSp"
repeat_region 19497..23797
/rpt_family="L1PAS"
SRS 22116..22396
/standard_name="DLIS3206"
SRS 22258..22436
/standard_name="DLIS2442"
repeat_region complement(24720..24929)
/rpt_family="MER30"
repeat_region complement(25741..26040)
/rpt_family="AluSx"
```

```
Query Match 3.9%; Score 317; DB 9; Length 94158;
Best Local Similarity 75.0%; Pred. No. 1,1e-55;
Matches 482; Conservative 0; Mismatches 140; Indels 21; Gaps 6;

QY 1175 GGGGGAAGAAAGCTGTGATACATTTCTTCTCCATCATTAAGGGTCGACCAAGACCTTA 1234
DB 72007 GGGCGAAGAAAGTAATGATACCTTCTCCGCCATCATTAAGGGTCATGGCTGACATTCCTA 71948
QY 1235 TAGTAAAGACAGGTATTAATAGACCAAAACCTAACAAATTATTATTATCAAACTTTTAA 1294
DB 71947 TAACCAAGATAGGTATTAATAGCAAAAC--ATAATGTATTATCAACTTTTATA 71891
QY 1295 TGACATGGAGCTTTTCAGAAATGAAGCCCAAGACCCAGGSAACCTGTCTTTT 1354
DB 71890 TTACATGGGAACCTTCATTAATGAAGACCAAGACCCAGGSAATCTGTCT-ATTTTTA 71832
QY 1355 TGCTAGGCTCGATGAAGAAATGATAGCATGTAGCATGTAGATTAGACAAAGATATG 1414
DB 71831 TGATTAGCTTGACGAAGAAATAGAGATCATGTAGCAATGTATGACAAAGGCTATG 71772
QY 1445 ATCTAGTGTAAGGACTCAGGGGAAACACAGCAGGCTGTCTATTCAATTTCTTT 1474
DB 71771 ATCTAATGTATAGACTAAGGGGAAAAACCCACAGGCTGTCTGTCCAGATACCTCTT 71712
QY 1475 GATCTCTCTCTGTATGATATGACATTCCTTCCCTCGAGTATGGGACGACTTCTT 1534
DB 71711 GACTGTCTC-----TGAGCTTCTCTTCTTCTTGAGATATGGGACGACTCTCCA 71660
QY 1535 CAATGAGGCTTTCAGGAGAGAGAGAGAAAGTGCCCTTTTATGATTTATGCTTGCTT 1594
DB 71659 AATATGAGCTTTCAGGAGAGAGAGAGAGAGAGACGTCTTCATGATTTATGCTTGCA 71600
QY 1595 CCGGGAAGAGAGATTCATGATTTCTATGACCATCTTGGGGAAGAGATTCGTTCTG 1654
DB 71599 TGTGGAGATGGTTCGTGTCTCAGACCTGTCTTGAGAAAGAAATCTGTCTTA 71540
QY 1655 TGACTTCTTTCATGAAGAAAGAGAGATTAAGAGCAGAGGAGAGATGTGCAGAAAG 1714
DB 71539 TGACTCTCTTGGGACAGAAAGAGAGCAGACAGCAAGAGGACAGATTAAGTCAGAG 71480
QY 1715 AGACTTGCCTGCTTCTAGG-GCTTCCGCTCTCTTTAGTTCCAAGTACTTCTTACATA 1773
DB 71479 A-----CCTGCTCTAGAGGCTCCAAATCTTCTTAAGTCAAGTA--CTCAGCATG 71428
QY 1774 CCAGGACATTAACCTTGGCATATGTTTCTGAGCTTACA 1816
DB 71427 CCAGGACCATTAACCTTGGATATCAAGTCTGAGCTCAACA 71385

RESULT 8
AC137626/c 175440 bp DNA linear HTG 27-MAR-2003
LOCUS Homo sapiens clone CTD-2145022, WORKING DRAFT SEQUENCE, 19
DEFINITION unorderd pieces.
ACCESSION AC137626.2 GI:29293979
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 175440)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-ouman,F.R., Allen,C.,
Alshrocks,S.L., Amaralunge,H.C., Are,U.R., Ayale,M., Banks,T.,
Barbasta,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieve,M., Brown,M., Brown,M., Bryant,N.P.,
Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.W.D., Dabchone,S.R., David,R.,
Dayala,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinb,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
```



Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabris, A., Gao, U., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J., H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlik, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hughes, M., Hollway, C., Hollins, B., Homel, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L.E., Jackson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kravovic, A., Kuresh, A., Landry, N., Leal, B., Lewis, J.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, U., Liu, W., Louised, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Madeshwar, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, A., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokkwo, S., Ogun, M., Okunolu, G., Ogunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shochetari, N., Sisson, I., Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Swalek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Taney, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Umami, K., Vasquez, U., Vera, V., Villalob, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallingford, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. P., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 175440)  
Worley, K.C.  
Submitted (27-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 175440)  
Worley, K.C.  
Direct Submission  
Submitted (27-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Mar 27, 2003 this sequence version replaced gi:25696694.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Drafting Center Code: BCM  
Contact: hgsc-help@bcm.tmc.edu  
Project Information  
Center project name: HERR  
Center clone name: CTD-2145022  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 172016 bases at least Q40  
Consensus quality: 174734 bases at least Q30  
Consensus quality: 176344 bases at least Q20  
Estimated insert size: 168157; sum-of-coverage estimation  
Quality coverage: 14x in Q20 bases; sum-of-coverage estimation

----- NOTE: Estimated insert size may differ from sequence length  
(see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 19 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1362: contig of 1362 bp in length  
\* 1363 1462: gap of unknown length  
\* 1463 3030: contig of 1568 bp in length  
\* 3031 3130: gap of unknown length  
\* 3131 4717: contig of 1587 bp in length  
\* 4718 4817: gap of unknown length  
\* 4818 6158: contig of 1341 bp in length  
\* 6159 6258: gap of unknown length  
\* 6259 7733: contig of 1475 bp in length  
\* 7734 7834: gap of unknown length  
\* 7834 9184: contig of 1351 bp in length  
\* 9185 9284: gap of unknown length  
\* 9285 11256: contig of 1972 bp in length  
\* 11257 11357: gap of unknown length  
\* 11357 12605: contig of 1250 bp in length  
\* 12607 12707: gap of unknown length  
\* 12707 14194: contig of 1488 bp in length  
\* 14195 14294: gap of unknown length  
\* 14295 15764: contig of 1470 bp in length  
\* 15765 15864: gap of unknown length  
\* 15865 17682: contig of 1818 bp in length  
\* 17683 17782: gap of unknown length  
\* 17782 18813: contig of 1031 bp in length  
\* 18814 18913: gap of unknown length  
\* 18914 20063: contig of 1150 bp in length  
\* 20064 20163: gap of unknown length  
\* 20164 21187: contig of 1024 bp in length  
\* 21188 22341: gap of unknown length  
\* 22341 22442: contig of 1054 bp in length  
\* 22442 22442: gap of unknown length  
\* 22442 23780: contig of 1339 bp in length  
\* 23781 23880: gap of unknown length  
\* 23881 25373: contig of 1493 bp in length  
\* 25374 25474: gap of unknown length  
\* 25474 27379: contig of 1906 bp in length  
\* 27380 27479: gap of unknown length  
\* 27480 175440: contig of 147961 bp in length.

FEATURES  
source  
1. 175440  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="CTD-2145022"

ORIGIN  
Query Match 3.9%; Score 317; DB 2; Length 175440;  
Best Local Similarity 75.0%; Pred. No. 1e-55;  
Matches 482; Conservative 0; Mismatches 140; Indels 21; Gaps 6;  
1175 GGGGTGAAACGCTGATTAATTTCTTCCTTCATCAAGGTCGCAACCAACTCTTA 1234  
114048 GGGCGAAGAGTAATGATACCTTTCTCGCCATCATAGGTCATGCGATTCCTTA 113989  
1235 TAGTAAAGACAGGTATATAAGACAAACCTTAACAATTATTTATCAAGTTTACA 1294  
113988 TAACCAAGATAGATTATATAAGAAAAC--ATAATGTATTTATCAACATTTATA 113932  
1235 TGACATGGGATCTTTCAGAAATGAGACCAAGACCAAGGAACTGCTTTTCTT 1354  
113931 TTACATGGGACCTTCATTATGAGACCAAGACCAAGGAAATGCTCT ATTITTA 113873  
1355 TGCTGAGGTGATGAAGATGATAGCATGTAGCCATGTAGATTAGACAAAGATATG 1414  
113872 TGATTAGGTTGAGAGGAATAGAGAGTCATGTAGCATGTATGACAAAAGGATG 113813  
1415 ATCTAGTGTAAAGACTCAGGGGAAACAGACAGAGCTGTCTATTCAGATTTCTTT 1474  
113812 ATCTAATGTATATGACTAAGGAAAAACACACAGGCTGTCTGTCCAGATCTTCTT 113753  
1475 GATCTCTCTCTCTATGTATAGCATCTTCTCTCGATATGAGGAGAGACTTCTTT 1534  
113752 GACCTGTCTC-----TGTAAGCTTTCTCTTTCGATGATAGGAGAGACTCTCCA 113701

OY	1535	CATGAGGGGTTCACAGGAGAAGAACTGCCTTTTACTATTATGCGTGCCCTT	1594
Dd	113700	AATAAAGGCTTCAAGGAAGAAAGGACAETGACCCTTTTAGTTTTATGCGTTCAT	113641
OY	1595	CGGGAAGAAGCAGCTTTCTAGTTTCTATGACCCACTTTGGGGAAGGAATTCGGTTTCGTG	1654
Dd	113640	TGTGGAGTAAGGGTCTGTGGTTCTCATGACCTGTCTTGGAGAAGGAATCTGTTTTCTA	113581
OY	1655	TEACTTGCTTTCATGAAGAAAAGAGATPAAGGACAGAGGCCGAGATGTGTAGAAC	1714
Dd	113580	TGACTCTCTTCGGGACAAAAAGAGGACAGCACAGAAAGGACAGATAAAGTACAGAG	113521
OY	1715	AGACTGGGCTGCTTGTGAGG-GCTTCGCGCTCTCCTTTAGTTCAGACTATCTTAGACATA	1773
Dd	113520	A-----CCTTGCTCTGAGGTCTTCGCTTCATCTTTTAGTTCAAAGR---CTCAGCATG	113469
OY	1774	CCAAGACATATACCTTTGCGCATPAGTGTTCCTGAGCCTTAACA	1816
Dd	113468	CCAAGGACCATACCTTTGAGTATCATAGTCTGAGCCTTAACA	113426
RESULT 9			
LOCUS	AC073227/c	154125 bp DNA linear HTG 12-AUG-2000	
DEFINITION	Homo sapiens chromosome 11 clone RP11-652LJ3, WORKING DRAFT		
SEQUENCE	45 unordered pieces.		
ACCESSION	AC073227		
VERSION	AC073227.1 GI:8440039		
SOURCE	HTG; HTGS_PHASE1; HTGS_DRAFT.		
ORGANISM	Homo sapiens Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 154125) Waterston,R.H. The sequence of Homo sapiens clone Unpublished 2 (bases 1 to 154125) Waterston,R.H. Direct Submission Submitted (10-JUN-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA		
JOURNAL			
COMMENT			
	Center: Washington University Genome Sequencing Center		
	Center code: WDGSC		
	Web site:http://genome.wustl.edu/gsc/index.shtml		
	Project information ----- Center project name: H_NH0652LJ3		
	Summary Statistics ----- Sequencing vector: MJ3; 100%		
	Chemistry: Dye-primer ET; 100% of reads Chemistry: Dye-terminator Big-Dye; % of reads Assembly program: Phrap; version 0.990319		
	Consensus quality: 131306 bases at least Q40 Consensus quality: 137026 bases at least Q30 Consensus quality: 139583 bases at least Q20		
	Insert size: 124000; agarose-fp Insert size: 149725; sum-of-contigs Quality coverage: 3.61 in Q20 bases; agarose-fp Quality coverage: 3.15 in Q20 bases; sum-of-contigs		
* NOTE:	This is a "working draft" sequence. It currently consists of 45 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.		
*	1	1169: contig of 1169 bp in length	

\* 105193 109346: contig of 4154 bp in length  
\* 109347 109346: gap of unknown length  
\* 109447 115145: contig of 5699 bp in length  
\* 115146 115245: gap of unknown length  
\* 115246 121014: contig of 5769 bp in length  
\* 121015 121114: gap of unknown length  
\* 121115 126886: contig of 5772 bp in length  
\* 126887 126886: gap of unknown length  
\* 126987 132609: contig of 5623 bp in length  
\* 132610 132709: gap of unknown length  
\* 132710 138881: contig of 6172 bp in length  
\* 138882 138981: gap of unknown length  
\* 138982 146419: contig of 7438 bp in length  
\* 146420 146519: gap of unknown length  
\* 146520 154125: contig of 7606 bp in length.  
Location/Qualifiers

FEATURES  
source  
1. .154125  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/clone="RP11-652J13"

misc\_feature 1. .1169  
/note="assembly\_name:Contig11"  
misc\_feature 1270. .2769  
/note="assembly\_name:Contig15"  
misc\_feature 2870. .4193  
/note="assembly\_name:Contig16"  
misc\_feature 4294.5710  
/note="assembly\_name:Contig18"  
misc\_feature 5811. .6984  
/note="assembly\_name:Contig19"  
misc\_feature 7085. .8592  
/note="assembly\_name:Contig20"  
misc\_feature 8693. .9797  
/note="assembly\_name:Contig21"  
misc\_feature 9898. .11428  
/note="assembly\_name:Contig22"  
misc\_feature 11529. .12866  
/note="assembly\_name:Contig23"  
misc\_feature 12967. .14965  
/note="assembly\_name:Contig24"  
misc\_feature 15066. .16845  
/note="assembly\_name:Contig25"  
misc\_feature 16946. .19643  
/note="assembly\_name:Contig26"  
misc\_feature 19744. .21949  
/note="assembly\_name:Contig27"  
misc\_feature 22050. .24411  
/note="assembly\_name:Contig28"  
misc\_feature 24512. .27410  
/note="assembly\_name:Contig29"  
misc\_feature 27511. .30054  
/note="assembly\_name:Contig30"  
misc\_feature 30155. .32447  
/note="assembly\_name:Contig31"  
misc\_feature 32548. .34490  
/note="assembly\_name:Contig32"  
misc\_feature 34591. .36874  
/note="assembly\_name:Contig33"  
misc\_feature 36975. .39829  
/note="assembly\_name:Contig34"  
misc\_feature 39930. .42666  
/note="assembly\_name:Contig35"  
misc\_feature 42767. .46079  
/note="assembly\_name:Contig36"  
misc\_feature 46180. .49377  
/note="assembly\_name:Contig37"  
misc\_feature 49478. .52066  
/note="assembly\_name:Contig38"  
misc\_feature 52167. .55853  
/note="assembly\_name:Contig39"  
misc\_feature 55954. .58804

/note="assembly\_name:Contig40"  
58905. .62776  
/note="assembly\_name:Contig41"  
misc\_feature 62877. .66204  
Query Match 3.9% Score 311.2; DB 2; Length 154125;  
Best Local Similarity 70.5%; Pred. No. 1,7e-54;  
Matches 509; Conservative 0; Mismatches 186; Indels 25; Gaps 6;

QY 1177 GGTGAAACGTGTATACATTTCTCTTCATATTAAGGTCGCAACCAACCTCTAT 1236  
DB 73372 GGGTAAAGAGTGTGTGCTTTCTTCCTCATTAAGGCTCAGCTGACCTCTATC 73313  
QY 1237 GTAAAGACAGCTTAATAAGACCAAACTAACAAATTTATTAATCAAGTTTACAG 1296  
DB 73312 ATAAAGACAGCTTAACAAGACCAATTAATTAATTAATTAATTAATTAATTA 73253  
QY 1297 ACATGGAGTCTTCAAAATGACCAACCAAGCCAGGGGAACTGTCTGTTTTTTT 1356  
DB 73252 ATACAGAGCTTCAAAATGACCAATTAATTAATTAATTAATTAATTAATTA 73194  
QY 1357 CTGAGTTCATGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1416  
DB 73193 CTGAGTTCATGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 73135  
QY 1417 CTAGTGTAAAGACTCAGAGGGAACACAGCAAGGCTGTCTATTAGATTTCTTGA 1476  
DB 73134 CTATGATTAAGACT----GAGAAACCAAGCAAGGCTGTATGTCAGATTTCTTCGG 73079  
QY 1477 TCTCTCTCTCTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1536  
DB 73078 TCTCTC-----TGATAGCAATTCCTCCTCTAGTATAGGAGAAACCCCTGTGA 73027  
QY 1537 ATGAGGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1596  
DB 73026 ATGAGGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 72967  
QY 1597 GGGAGAGGAGTCTTATGTTCTATGACCATCTTGGGAGAGGAATCTGTTCTG 1656  
DB 72966 GGGAGAGGAGTCTTATGTTCTATGACCATCTTGGGAGAGGAATCTGTTCTATG 72907  
QY 1657 ACTTGCTTCTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1716  
DB 72906 ACTCATTTAGGGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 72847  
QY 1717 ACTTGCTGCT-----TCTGAGGCTTCGCTCTCTTATGTTCAAGTACTTCTTA 1768  
DB 72846 ACTTGCTTCTGAGGCTACTTCTGAGCTTCAATCTCTTGTGTTCAAGTACTTCTTA 72790  
QY 1769 GCATACCAAGCATATATCTTTGGCATATGTTTCTGAGCTTAACTGCAATATGC 1828  
DB 72789 GCATGCCAAAGTCAATATCTTGCATATCTTCTGAGCCCAAGAGAGAAAGCTA 72730  
QY 1829 TAACTCCTCATGACCTCAACATTCACCTTGCTTTTATCTTTATGTTGTGATGC 1888  
DB 72729 AAATGCTCATATACCAACCAACCAACCAATTAAGTTTATACATTCATAGAGAT 72670  
QY 1889 AT 1890  
DB 72669 AT 72668

RESULT 10  
AC018694 196832 bp DNA linear PRI 09-MAY-2001  
LOCUS AC018694  
DEFINITION Homo sapiens BAC clone RP11-563P16 from 11, complete sequence.  
ACCESSION AC018694  
VERSION AC018694.4 GI:9665213  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 196832)  
AUTHORS Sulston, J.E. and Waterston, R.  
TITLE Toward a complete human genome sequence  
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)  
MEDLINE 99063792  
PUBMED 9847074

REFERENCE 2 (bases 1 to 196832)  
AUTHORS Ali, J., Maupin, R. and Lehnert, U.  
TITLE The sequence of Homo sapiens BAC clone RP11-563P16  
JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 196832)  
AUTHORS Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (16-DEC-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 4 (bases 1 to 196832)  
AUTHORS Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (03-AUG-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 5 (bases 1 to 196832)  
AUTHORS Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (10-SEP-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 6 (bases 1 to 196832)  
AUTHORS Waterston, R.  
TITLE Direct Submission  
JOURNAL Submitted (08-NOV-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

REFERENCE 7 (bases 1 to 196832)  
AUTHORS Waterston, R.  
TITLE Direct Submission  
JOURNAL Submitted (09-MAY-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

COMMENT On Aug 3, 2000 this sequence version replaced gi:7631073.  
----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc>  
Contact: [sapiens@wustl.edu](mailto:sapiens@wustl.edu)  
----- Summary Statistics  
Center project name: H\_NH0563P16  
-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:  
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frenken, R., Tatenno, M., Caranese, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from

Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)  
VECTOR: pBAC3.6  
NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the left is RP11-484L12. The clone sequenced to the right is RP11-652L13. Actual start of this clone is at base position 1 of RP11-563P16; actual end is at base position 196832 of RP11-563P16.

The region from base position 99028 to 99031 is covered by a single m13 subclone.

FEATURES  
source  
location/Qualifiers  
1..196832  
/organism="Homo sapiens"  
/mol\_type="Genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11"  
/clone="RP11-563P16"  
/clone\_1fb="RPCI-11"  
25..466  
/rpc\_family="L1"  
repeat\_region  
917..1135  
/rpc\_family="L2"  
repeat\_region  
1137..1361  
/rpc\_family="MIR"  
repeat\_region  
1386..2075  
/rpc\_family="L2"  
repeat\_region  
2132..4576  
/rpc\_family="L1"  
repeat\_region  
4616..4774  
/rpc\_family="L2"  
repeat\_region  
5092..5202  
/rpc\_family="MIR"  
repeat\_region  
5241..5572  
/rpc\_family="ERV"  
repeat\_region  
5884..6021  
/rpc\_family="L2"  
repeat\_region  
6078..6151  
/rpc\_family="L1"  
repeat\_region  
6301..6477  
/rpc\_family="MER1\_type"  
repeat\_region  
6597..6704  
/rpc\_family="L1"  
repeat\_region  
8852..8954  
/rpc\_family="MIR"  
repeat\_region  
8954..9034  
/rpc\_family="L2"  
repeat\_region  
9411..9560  
/rpc\_family="ALU"  
repeat\_region  
10259..10313  
/rpc\_family="L2"  
repeat\_region  
10395..10599  
/rpc\_family="MIR"  
repeat\_region  
10743..11056  
/rpc\_family="ALU"  
repeat\_region  
11522..11673  
/rpc\_family="L1"  
repeat\_region  
13764..13928  
/rpc\_family="L2"  
repeat\_region  
14035..14188  
/rpc\_family="L2"  
repeat\_region  
14229..14659  
/rpc\_family="L1"  
repeat\_region  
14660..14751  
/rpc\_family="L2"  
repeat\_region  
14802..15365  
/rpc\_family="MER2\_type"  
repeat\_region  
16007..16389  
/rpc\_family="MER1\_type"  
repeat\_region  
17025..17123  
/rpc\_family="ERV1"

```
misc_feature 17033..17293 EST BE005163 (NID:g8265396)
misc_feature 17050..17293 EST AA528257 (NID:g2270326) n26a02.s1"
misc_feature 17053..17293 EST A1350387 (NID:g4087593) q039d10.x1"
misc_feature 17054..17293 EST N57929 (NID:g1201819) yv61e06.s1"
repeat_region 17648..17735 /rpt_family="MIR"
repeat_region 19011..19327 /rpt_family="MIR"
repeat_region 21465..21556 /rpt_family="MIR"
repeat_region 21718..21927 /rpt_family="Alu"
repeat_region 21961..22078 /rpt_family="MIR"
repeat_region 22071..22284 /rpt_family="MER1_type"
repeat_region 22285..22589 /rpt_family="MER1_type"
repeat_region 22590..22729 /rpt_family="Alu"
repeat_region 22745..22943 /rpt_family="MER1_type"
repeat_region 23885..23989 /rpt_family="MIR"
repeat_region 24058..24464 /rpt_family="MIR"
repeat_region 24466..24582 /rpt_family="L1"
repeat_region 24583..24880 /rpt_family="L1"
repeat_region 24881..25141 /rpt_family="Alu"
repeat_region 25149..25336 /rpt_family="L1"
repeat_region 25948..26580 /rpt_family="MALR"
repeat_region 27227..27361 /rpt_family="ERV1"
repeat_region 28278..28567 /rpt_family="MIR"
repeat_region 28661..29094 /rpt_family="Alu"

Query Match 3.9% Score 311.2; DB 9; Length 196832;
Best Local Similarity 70.5%; Pred. No. 1.7e-54;
Matches 509; Conservative 0; Mismatches 188; Indels 25; Gaps 6;

1177 GGTGAAAAGCTGTATACATTTCTCTTCATCATAGGGGTGCGAACAAACTCTTATA 1236
25963 GGGTGAAGAGTGTGTACCTTTCTTGCCCATCATAGGGGTGCGACGCTGCGACCTCATC 26012

QY 1237 GTAAAAAGCAGTTATATAGACAAACCTTAAACAATTTATTATCAAAAGTTTACATG 1296
DB 26013 ATAAAAAGCAGTTATATAGACAAACCTTAAACAATTTATTATCAAAAGTTTGTGCG 26072
QY 1297 ACATGGAGTCTTCGAAATGAGACCAAGACCCAGGGGAAACTGTCTGTTTTTTT 1356
DB 26073 AATACGAGAGCCTTCGAATTAAGTCAAGAACTTGAGAGAAACCTCTG-CTTTATG 26131
QY 1357 CTGAGGTTGATGAAGATGATAGCATGTAGCCATGTAGATTAGACAAAGATATGAT 1416
DB 26132 CTTAGTTTCAAGGAATGACCAACCATGTAGCAATGTGATT-GACAAAGGCTTGAT 26190
QY 1417 CTAGTGGTAAAGCATCGGGGAAACACAGACAGGCGCTGTCTATTGAGATTCCTCTGA 1476
DB 26191 CTAAATGATATGACT---GAGAAACCAACAGGCTGTGATTCAGATTCCTTCGG 26246
QY 1477 TCTCTCTCTCTATGTATAGCATTCCTCTCTCTGATATGGGCGAGACTCTTCTCA 1536
DB 26247 TCTCTC-----TGATAGCATTCCTCTCTCTAGTATAGGCGAGAACCCCTGTGGA 26298
```

```
QY 1537 ATGAGGCTCTCAAGGAGAGAGGAGAAAGTGCCTTTTATGATTTATGCTTGTCTCG 1596
DB 26299 ATGAGGCTCTCAAGGAGAGAGGAGAAAGTGCCTTTTATGATTTATGCTTGTCTCG 26358
QY 1597 GGGAGAGAGAGTTTATGATTTCTATGACCCATCTTGGGGAAGAGAAATTCGTTTCTG 1656
DB 26359 GGGAGAGAGAGTTTATGATTTCTATGACCCATCTTGGGGAAGAGAAATTCGTTTCTG 26418
QY 1657 ACTTGCTTTCATGAAGAAGAGAGATGAGAGGAGGAGGAGAGATGCTGAGAAAG 1716
DB 26419 ACTGACTTTAGGGAGATGAGAGGAGGAGAAATGAGAGGCCAGAGAGATGAGAGAG 26478
QY 1717 ACTTGCTGCT-----TCTGAGGCTTCGCGCTCTCTTTCCTTTCAGTTCCTT 1768
DB 26479 ACTTGTCTTCGAGGCTACTCTTGAGCTTCGACCTTCATCTCTTTCCTTTCAGTA 26535
QY 1769 GCATACCAAGACACTATGCTTGGCATATGTTTCTGAGCTTAAACATGCAATCATG 1828
DB 26536 GCATGCCAAATGTCATCTTTCGAGTATATTTTCGAGCCCAAGAGAGAAAGCTA 26595
QY 1829 TAACTCCTCTATACCTTCAACATTCGACTGCTTTTATCTTATGTTGTGTGATGC 1888
DB 26596 AAATGGCTCATATATCCACCAACGACATTAAGGTTTATCATTCAGTTGAGAGAT 26655
QY 1889 AT 1890
DB 26656 AT 26657

RESULT 11
AC009867 176773 bp DNA linear HTG 22-NOV-2000
LOCUS Homo sapiens chromosome 11 clone RP11-76K11 map 11, WORKING DRAFT
DEFINITION SEQUENCE, 12 unordered pieces.
ACCESSION AC009867
VERSION AC009867.2 GI:11276242
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 176773)
Biren,B., Binton,L., Nusbaum,C., and Landry,E.
Homo sapiens chromosome 11, clone RP11-76K11
unpublished
2 (bases 1 to 176773)
Biren,B., Binton,L., Nusbaum,C., Landry,E., Allen,N., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckert,R., Benn,U., Brown,A.,
Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeRellano,K., Depayre,E., Devon,K., Dewar,K.,
Donelan,L., Doyle,M., Ferreira,P., Fitzhugh,N., Forrest,C.,
Funke,R., Gage,D., Galagan,J., Gardina,S., Gilbert,D., Grant,G.,
Hagos,B., Hesford,A., Horton,L., Howland,J.C., Jones,C., Kam,L.,
Karatas,A., Lehotzky,D., Lien,C., Locke,K., Macdonald,P.,
Marquis,N., McSwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
Meldrum,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Naylor,J., Nijoff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
Stefange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
Tessaye,S., Tortorella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
Wheeler,D., Wu,X., Wyman,D., Ye,W.U. and Zody,M.
Direct Submission
Submitted (03-SEP-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 22, 2000 this sequence version replaced gi:5822691.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
```





QY	1221	AACCAAAACCCCTATGTAAAGACAGGTTATTAAGACAAACCTTAACAAATTTATTTA	1280
Db	168488	GGCACACACTCTTATACAAAGGACAGAGAACAGAGAAATGACATGACAAATATTTTA	1684229
QY	1281	ATCAAAAGTTTACATGACATGGGAGCTTTCAGAAAATGAAGCCCAAGACCAGGGAAA	1340
Db	168428	ACCAAAAGTTTATGTGACATGGGAGCCTTGAAGAAATTAAGACCAAAAGCCACAGGAAAA	1683659
QY	1341	CTCTCGTTTTTTTTTCTGAGGTTGAGTGAAGAAATGATACATGTAGCCATGTAGTTA	1400
Db	168368	TTTCTCTG-TTTTATGCTTTAGGTTTGAATGAAGAACTGCACGCCATATAGAAATGTGATTA	1683104
QY	1401	GACAAAAGCATATGATCTAGTGTGAAGAGCTCA-GGGGAGAAACACAGCAAGGCTCTCT	1459
Db	168309	GACAAAAGGTATGATCTTAATCTTAACAGACTGACAGGAGAAACCCAGCCAGGCTCCCT	168250
QY	1460	ATTACAGATTTCTTTCATCTCTCTCTCTATATGTATGACATTTCTTCTCTAGTATGG	1519
Db	168249	TT-----CTGTTTCAGATTTCTTCTCGGCTCCCATATTAACATTCCTCTCTGGGTATGA	168194
QY	1520	GGCAGGACCTCTTCTTCATGAGAGCTTTCAGAGGAGAGGAGAGAAAGTGGCTTTTAGA	1579
Db	168193	GGCTGACCTCTCTGGAATGAGATTTCTTCAGACAGAAAGAGAGATGACCTTCTTAGG	1681944
QY	1580	TTTTATGCTTCTGCTGGGAGAAAGAGAGTTCTAGTTTATGACCCATCTTGGGAGAG	1639
Db	168133	TTTTATGCGATCTGCTGAGAGAAAGAGAGTTCTGTTTCTATGACCCACTTGGGAGAG	168074
QY	1640	GAATTCGTTGTTCTGTGACTTCTTCATGAAAGAGAGAGCTAAGAGGCGAGGGGAG	1699
Db	168073	GAGTTCTGGTTTATATGACTCACTTAATCTTAGCCGAGAAAGAGACAGAGATAGAGG	168014
QY	1700	GAGATGATCAGAAAGAGACTGCTGCTCTTCTAGAGGCTTCGCTCTCTTTAGTTCCAG	1759
Db	168013	GTAAGAAAGAGTCAAGAAACCTGTTTCTGAGTCTTCCAAATGTCCTTCAAGTTCAAAG	1679544
QY	1760	TACTTCTTAGCATACCAAGACACTATCTTTTGGCATATGTCTTCTGAGCTCTAACCTG	1819
Db	167953	TA---CTCAGCATGCCAGAGTCCATCTTTGGGGTATCAGATTTCTGAGCCCCAAACCTC	1678979
QY	1820	CAA 1822	
Db	167896	CTA 167894	
RESULT 14			
	AL591491	121720 bp	DNA linear PRI 06-DEC-2001
LOCUS	AL591491	Human DNA sequence from clone RP11-435K1 on chromosome Xp15.3-15.5	
DEFINITION	AL591491	Containing the 5' end of a novel gene, complete sequence.	
ACCESSION	AL591491.5	GI:15149577	
VERSION	AL591491.5		
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Ekayachia, Merzosa; Chordara; Cranlara; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 121720)		
TITLE	Lawlor, S.		
JOURNAL	Direct Submission		
COMMENT	Submitted (06-DEC-2001) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: clonerequest@sanger.ac.uk humbrey@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Aug 13, 2001 this sequence version replaced gi:15020561. Where difference assembly data is compared from overlapping clones. During difference assemblies are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >=		

## FEATURES

```

30) ; an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep/ This sequence
was generated from part of bacterial clone contigs of human
Chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX
Rpl1-435K1 is from the library RPCI-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6
IMPORTANT: This sequence is not the entire insert of clone
Rpl1-435K1. It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true left end of clone Rpl1-435K1 is at 1 in this sequence. The
true left end of clone Rpl1-342D14 is at 119721 in this sequence.
UNRES
Location/Qualifiers
1..121720
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="p15.3-15.5"
/clone="Rpl1-435K1"
/clone_lib="RPCI-11.2"
17..90
/note="MIR repeat: matches 28..111 of consensus"
repeat_region
443..559
/note="MERSA repeat: matches 59..189 of consensus"
repeat_region
1986..2375
/note="match: GSS: Em:A0725819"
misc_feature
2022..2248
/note="match: GSS: Em:A0621701"
repeat_region
2544..2635
/note="46 copies 2 mer tt 73% conserved"
2638..2932
/note="Alu10 repeat: matches 1..312 of consensus"
repeat_region
3343..3705
/note="12 repeat: matches 2370..2746 of consensus"
repeat_region
4022..4217
/note="MIR repeat: matches 8..208 of consensus"
4624..4797
/note="87 copies 2 mer ta 69% conserved"
5082..5534
/note="match: GSS: Em:A0442500"
misc_feature
5095..5533
/note="match: GSS: Em:A0442489"
6039..6101
/note="MSTB repeat: matches 1..64 of consensus"
repeat_region
6102..6395
/note="Alu10 repeat: matches 25..309 of consensus"
repeat_region
6366..6768
/note="MSTB repeat: matches 64..426 of consensus"
6979..7100
/note="12 repeat: matches 2575..2709 of consensus"
repeat_region
8039..8302
/note="MER20 repeat: matches 1..216 of consensus"
repeat_region
8324..8823
/note="LINE repeat: matches 943..1433 of consensus"
8829..9032
/note="LINE3 repeat: matches 3815..4011 of consensus"
repeat_region
9478..9746
/note="LINE3 repeat: matches 4058..4325 of consensus"
9747..10084
/note="MER44A repeat: matches 1..332 of consensus"
10085..10943
/note="LINE3 repeat: matches 4325..5186 of consensus"

```



```
repeat_region 10947..11151
/note="AluJo repeat: matches 86..297 of consensus"
repeat_region 11153..11917
/note="MER3 repeat: matches 5182..6002 of consensus"
repeat_region 11918..12011
/note="MER81 repeat: matches 2..114 of consensus"
repeat_region 12012..12175
/note="LIMB3 repeat: matches 6002..6156 of consensus"
repeat_region 14703..15013
/note="AluY repeat: matches 1..111 of consensus"
repeat_region 15834..15923
/note="L2 repeat: matches 2574..2666 of consensus"
repeat_region 16531..16566
/note="18 copies 2 mer ca 100% conserved"
repeat_region 17622..18127
/note="LIMB5 repeat: matches 5679..6174 of consensus"
repeat_region 18145..18248
/note="FLAM C repeat: matches 5..108 of consensus"
repeat_region 18392..18655
/note="AluSG repeat: matches 1..305 of consensus"
repeat_region 19551..19659
/note="L2 repeat: matches 2549..2662 of consensus"
gene complement(20811..29961)
/gene="DA435K1.1"
complement(join(<20811..21075,29654..29961))
/product="DA435K1.1 (novel protein)"
/note="match: CDNAS: Em:AK024165 Em:AK016298
match: ESTs: Em:AA420823 Em:AA455585 Em:AI796127
Em:AUI21832"
evidence=not experimental
complement(join(<20811..21075,29654..29886))
/gene="DA435K1.1"
/codon_start=1
/evidence=not experimental
/product="DA435K1.1 (novel protein)"
/protein_id="CAD13520.1"
/db_xref="GI:17426506"
/db_xref="SPTREMBL:O8WX00"
/translacion="MEPOLGEAALRPGWLALLIWLVSALSCFSI.PASSLSIVPOV
RYSNFGRTIGLDKCNACIGTSCCKCFEBEIRSDNWLASHGLPDDLSIPANTY
DPSKIRPVELFRUVSKRQNEISDRICASAPKTCISIRVLRKTERFOKMLQARLT
PDIVQ"
repeat_region 21413..21515
/note="MIR repeat: matches 42..147 of consensus"
repeat_region 21840..22137
/note="AluSX repeat: matches 2..299 of consensus"
repeat_region 23050..23260
/note="MIR repeat: matches 47..262 of consensus"
repeat_region 23316..23609
/note="AluSX repeat: matches 1..294 of consensus"
repeat_region 24752..25050
/note="L2 repeat: matches 2245..2543 of consensus"
misc_feature complement(25215..25383)
/gene="DA435K1.1"
/note="match: GSS: Em:AL013940"
complement(25215..25382)
/gene="DA435K1.1"
/note="match: GSS: Em:AL013938"
repeat_region 26303..26595
/note="AluSG repeat: matches 1..293 of consensus"
repeat_region 26651..26750
/note="LIM4 repeat: matches 5302..5399 of consensus"
repeat_region 29031..29118
/note="MIR repeat: matches 25..114 of consensus"
repeat_region 31059..31310
/note="LIM4 repeat: matches 5099..5385 of consensus"
repeat_region 31315..31499
/note="MER93 repeat: matches 205..397 of consensus"
repeat_region 31565..31643
/note="MER93 repeat: matches 3..82 of consensus"
repeat_region 31649..31998
/note="MER57-internal repeat: matches 6808..7145 of
```

```
consensus"
repeat_region 31998..32402
/note="MER65-internal repeat: matches 3648..4734 of
consensus"
repeat_region 32415..33449
/note="MER31-internal repeat: matches 1721..2860 of
consensus"
repeat_region 33450..33759
/note="AluSX repeat: matches 1..309 of consensus"
repeat_region 33868..35557
/note="LIM2 repeat: matches 4674..6306 of consensus"
repeat_region 35564..37204
/note="LIM1 repeat: matches 830..2496 of consensus"
repeat_region 37217..37364
/note="LIM1 repeat: matches 1264..1116 of consensus"
repeat_region 37382..39552
/note="LIM1 repeat: matches 1389..828 of consensus"
repeat_region 39638..39764
/note="MER34 repeat: matches 418..543 of consensus"
repeat_region 40368..40646
/note="MER31-internal repeat: matches 1117..1393 of
consensus"
repeat_region 41671..41824
/note="MER93 repeat: matches 248..394 of consensus"
misc_feature 41957..42120
/note="match: GSS: Em:AQ170009"
repeat_region 42164..42825
/note="LIM4 repeat: matches 4435..5094 of consensus"

Query Match 3.7%; Score 297.6; DB 9; Length 121720;
Best Local Similarity 68.6%; Pred. No. 1.2e-51;
Matches 495; Conservative 0; Mismatches 199; Indels 28; Gaps 5;

QY 1113 TAGGGGTGAAACCTGTGATACATTTCTCTCCATCATATAGGTCGACCAAACTCC 1232
DB 102339 TATGATATGAAGAGGTGGGCTTCCTTCCTTACCACATTAGGTCATGCCACACCCC 102458
QY 1233 TATGTAAGACAGCTTAATTAAGACAAACCTAACAAATTATTATTAACAAAGTTTA 1292
DB 102459 TATAACAAGGAAGTTTACCAAGAAAAGCATTAACAATTATTGATCAAAAGTTTA 102518
QY 1293 CATACATGGGAGTCTTCAGAAAATGAAGACCCAAAGACCCAGGGGAAGTCTGTTTT 1352
DB 102519 TGTACACAGGAGCTTCAGAAATGAAGACCCAAAGCCAGGTCAA--TGATTTATTT 102576
QY 1353 TTTCGAGGTGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1412
DB 102577 TGTCTTAAGTTGATGAAGATGGGAGGTATATGAATGGGATTTGACAAAGGCTG 102636
QY 1413 TGATCTAGTGAAGACTCAGGGGAAAACAGCAGAGGCTGTCTATTTCAGATTCCTC 1472
DB 102637 TGATCTGTGGTAAATAACTGAGGAGAAACCAACAAGTGTGTGTGATTCCTCC 102696
QY 1473 TTGATCTCTCTCTCTATGTATGATGATTTTCTCTCTGAGTATGGGACAGATCTTC 1532
DB 102697 TTGGCACTC-----TGTGATGACATTCCTCTCTCCTAGGATATGGGACAGATCTTC 102748
QY 1533 TTCAATAGAGGTCTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1584
DB 102749 TGAATAGAGGTCTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 102808
QY 1585 TGCTTCTCTGGGGAAGAGAGTCTTATGTTTATGATCCCATCTTGGGGAAGAGAGAT 1644
DB 102809 TAGCTTCTTGGGGAAGTGAAGTCTTATGTTTATGATCCCATCTTGGGGAAGAGAGAT 102868
QY 1645 CTGCTTCTGGAAGTCTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1704
DB 102869 CTGCTTCTTATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 102926
QY 1705 GGTGAGAAAGAGCTTGGCTCTTGGAGGCTTCGCTCTCTTATGTTCAAGTACTT 1764
DB 102927 GGTGAGAGAGAGATTAATCTTGAAGA-----CTTCTATATCTGTGATGTCAGACTA 102978
```

QY 1765 CTTAGCATCAACAAGCACTATACCTTTGGCATATGGTTTCTGAGCTCTAACCTGCATC 1824  
DB 102979 CTCAGCATCTAAGGTCGGTATTTTGGGATATGTCTTAAGCCCAAGTATATTA 103038  
QY 1825 AAGCTAACTCTCTATGACCTTCAACATTCACCTTCTTATCTTTATGCTGTA 1884  
DB 103039 TGGCTAAGATTTTCTCTGCAAAAATGTTTATCTATTTTATTAATCAGATGTA 103098  
QY 1885 TG 1886  
DB 103099 TG 103100  
RESULT 15  
AC010244/c 134937 bp DNA linear PRI 15-AUG-2001  
LOCUS AC010244 Homo sapiens chromosome 5 clone CTC-364B20, complete sequence.  
DEFINITION AC010244  
AC010244  
VERSION AC010244.6 GI:15187208  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 134937)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 134937)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE 3 (bases 1 to 134937)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (26-JAN-2001) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
REFERENCE 4 (bases 1 to 134937)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (15-AUG-2001) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
COMMENT On Aug 15, 2001 this sequence version replaced gi:12545284.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 99.6% of Sequence;  
Estimated Total Number of Errors is 0.4.  
STS Content:  
SHGC-79386 G52178.  
Location/Qualifiers  
1. 134937  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="CTC-364B20"  
FEATURES  
source  
ORIGIN  
Query Match 3.7%; Score 293.8; DB 9; Length 134937;  
Best Local Similarity 72.8%; Pred. No. 7.2e-51;  
Matches 481; Conservative 0; Mismatches 157; Indels 23; Gaps 7;  
QY 1162 CAAAGCTGACGTGGGCAAAACGTGATACATTTCTTCATCATTAAGGTCGCA 1221  
DB 22463 CTATCCAGCTGTCAGGTAAAGGATGATATGTTTCTCACCCTATATAGAGTCACA 22404  
QY 1222 ACCAAAACTCTTATAGTAAAGACAGGTTTATTAAGACAAAACCTTAACAATTTATTTAA 1281  
DB 22403 GCTGACATTCCTATTAACAAAAGATAGGTTATTAAGAGAAAGCATTAACACTTTTATTTAA 22344

QY 1282 TCAAAATTATACATGACATGGAGCTCTTCGAAATGAAGCCCAAGACCCAGGGGAAC 1341  
DB 22343 TCAAAATTATCTGATATACAAAGCCCTTCGAAATGAAGCCCAAGACCCAGGAAAAAC 22284  
QY 1342 TGTCTGTTTTTTTGGTGGAGGTGATGAAGATGATAGCATGTAGCCATGTAGATTAG 1401  
DB 22283 TG-CCCTTTTATGCTTATGTTTATGTAAGAAATGACACCCAGGTGAAATGTGATTGG 22225  
QY 1402 ACAAAGGATATGATCTAGTGTATTAAGAGTCAAGGGG-AAACACAGCAAGGCTGTCTA 1460  
DB 22224 ACAAAGGATATTAATTTAAAGTGTAGAGATGAGGGGGAACCAAGCAAGGCTGTCTG 22165  
QY 1461 TTCAGATCTCTGATCTCTCTCTCTATGATATGATCTCTCTCTCTCTCTCTCTCTCT 1520  
DB 22164 TTCAGATCTCTCTGATCT 22113  
QY 1521 GCAGGACTCTTCTCAATGAGGCTCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1580  
DB 22112 GCAGGATTAATCTGAAATGAGGCTCTGAAAGGAGGAGGAGGAGGAGGAGGAGGAGG 22056  
QY 1581 TTTATGCTTCT 1640  
DB 22055 TTTATGCTGATCT 21996  
QY 1641 AATCTGCTTCT 1699  
DB 21995 AATTTTGTCTTCT 21936  
QY 1700 GAGATGTCAGAAAGAGACTTGGCTCTCTGAGGCTTCCGCTCTCTCTCTCTCTCTCT 1759  
DB 21935 GAGAAATCAACAGCA-----CTTTCCTTTGAGGCTTCCAAATCTCTCTCTCTCT 21885  
QY 1760 TACTCTTATGATACCAAGACATATCTTTGGATATGTTTCTGAGCTCTAAGCTG 1819  
DB 21884 AAGTATATGATATCCAAAGATATGTTCTTTGGCTATCTGATCTCTGAGCCCAACAG 21825  
QY 1820 C 1820  
DB 21824 C 21824

Search completed: September 11, 2004, 16:46:42  
Job time : 19875 secs